AFRL-RQ-WP-TR-2019-0218

HARNESSING PARAMETERIZATION FOR FAST AND RELIABLE NONCONVEX OPTIMIZATION

Benjamin Recht

University of California, Berkeley

Pablo Parrilo and Asuman Ozdaglar

Massachusetts Institute of Technology

OCTOBER 2019
Final Report

AIR FORCE RESEARCH LABORATORY
AEROSPACE SYSTEMS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7542
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RQ-WP-TR-2019-0218 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signatures on File//

PHILIP S. BERAN, PhD
Work Unit Manager and Technical Advisor
Design and Analysis Branch

//Signatures on File//

CHARLES TYLER
Chief, Design and Analysis Branch
Aerospace Vehicles Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
October 2019	Final	5 April 2018 – 4 October 2019

4. TITLE AND SUBTITLE
HARNESSING PARAMETERIZATION FOR FAST AND RELIABLE NONCONVEX OPTIMIZATION

5a. CONTRACT NUMBER
FA8650-18-2-7839

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61101E

6. AUTHOR(S)
Benjamin Recht (University of California, Berkeley)
Pablo Parrilo and Asuman Ozdaglar (Massachusetts Institute of Technology)

5d. PROJECT NUMBER
CCS-02 (DARPA)

5e. TASK NUMBER

5f. WORK UNIT NUMBER
Q1YS

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
253 Cory Hall
Berkeley, CA 94720-1770

Massachusetts Institute of Technology

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Aerospace Systems Directorate
Wright-Patterson Air Force Base, OH 45433-7542
Air Force Materiel Command
United States Air Force

Defense Advanced Research Projects Agency/Defense Sciences Office (DARPA/DSO)
3701 N. Fairfax Drive
Arlington, VA 22203

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RQVC

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RQ-WP-TR-2019-0218

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES
This report is the result of contracted fundamental research, which is deemed exempt from Public Affairs Office security and policy review in accordance with Deputy Assistant Secretary of the Air Force (Science, Technology, Engineering) (SAF/AQR) memorandum dated 10 Dec 08 and Air Force Research Laboratory Executive Director (AFRL/CA) policy clarification memorandum dated 16 Jan 09.

14. ABSTRACT
This project focused on developing novel understanding of large-scale, non-convex optimization problems by establishing robust notions of how the choice of parameterization affects the geometric and computational character of the optimization process. This understanding was used to create a methodological link between machine learning and optimal control, enabling a car to be successfully taught to drive around an unspecified track using vision-based control. Reparameterization also provided benefits for optimization of recurrent neural networks. Insights were gained into the construction of well-performing stable recurrent models for future used in machine learning.

15. SUBJECT TERMS
non-convex optimization, recurrent neural networks, adaptive control, autonomous systems, vison-based control, robust learning

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
Unclassified	Unclassified	Unclassified	SAR	15	Philip S. Beran
					19b. TELEPHONE NUMBER (Include Area Code) N/A

TABLE OF CONTENTS

Section **Page**

LIST OF FIGURES

1. EXECUTIVE SUMMARY

This project was one of several fundamental efforts funded by the DARPA Defense Sciences Office through the Lagrange program. The Lagrange program sought to develop new mathematical approaches to optimization problems in uncertain, dynamic, multiscale, and high-dimensional systems. Lagrange aimed to enable solutions for complex, realistic problems involving dynamic environments, rapidly changing requirements, and increasing or decreasing amounts of information.

The primary technical goal of this project towards the broader Lagrange objectives was to explore the use of reparameterization to re-cast non-convex problems in larger parameter spaces in a manner improving problem convexity. The understanding gained through this exploration was applied to topics in machine learning and optimal control.

This report overviews the work accomplished in the project and provides citation to the many publications authored by the combined team at the University of California, Berkeley, and the Massachusetts Institute of Technology. These publications are the comprehensive archive of research documentation and provide details of method development, theoretical findings including mathematical proofs, and numerical results of benchmark cases. One journal article was published over this 18-month project. The remaining publications are available on arXiv; some of these articles are in peer review.

2. INTRODUCTION

Convex optimization, the mathematical process of optimizing convex functions over convex sets, has become an extremely reliable and practical technology marked by efficient, polynomial-time algorithms. However, algorithms that are generally reliable and efficient for non-convex optimization do not exist. This gap in capability is important: non-convex optimization is beginning to enable powerful industrial applications in machine learning, operations, planning, and controls. The goal of this fundamental research was to increase the reliability and efficiency of non-convex optimization for large-scale problems representative of challenging environments, and to explore the controlled conditions under which this improvement could be achieved.

The observation that inspired this work was that non-convex problems had been successfully optimized in the past in certain settings were there was significant flexibility in how the problem was formulated. In this project, reparameterization was formally introduced as an expression of this desired flexibility, and was used to develop novel understandings of large-scale, non-convex problems and to establish robust notions of how the choice of parameterization affects the geometric and computational properties of these problems.

The project achieved success in three main areas, described in Sections 3-5. Section 3 describes new insights gained into the application of machine learning to optimal control of uncertain and complex systems, as enabled by what is herein referred to as reparameterization. Section 4 describes new insights into recurrent neural networks and how reparameterization improves optimization under certain conditions in these models. In Section 5, the use of reparameterization to simplify the analysis and deployment of large-scale nonconvex optimization solvers is discussed.

3. MANAGING UNCERTAINTY AND COMPLEXITY BY MERGING MACHINE LEARNING AND OPTIMAL CONTROL

The first thrust of this investigation was the examination of how enlarging parameterization, referred to here as "reparameterization" or "lifting", simplifies non-convex optimization in machine learning and control. Of particular interest were efficient reparametrizations/liftings. One attractive method was the System Level Synthesis (SLS) parameterization [12], which models dynamical systems by their mapping from disturbance to state, rather than as a standard dynamical system. In this project, SLS was applied to controls; the resulting theoretical framework ultimately provided the first end-to-end bounds on the performance of a linear quadratic regulator when the dynamics are learned coarsely (i.e., the model is estimated from a few experimental trials) [3]. The key lesson was that all of this learning can be done with convex optimization by lifting the initial non-convex problem into a higher dimensional convex parameter space.

The research team also studied how to use stable models and concepts from the modern theory of robust control to enhance the stability of existing adaptive control algorithms for the control of autonomous vehicles (as described below, a car was algorithmically taught to drive a track autonomously). In adaptive control, an autonomous system begins operation while simultaneously learning about its dynamics and environment. The problem of adaptively controlling an unknown dynamical system has a rich history; classic asymptotic results of convergence and stability for adaptive control theory were already well established in the 1970s. Since this time, there have been several works that analyze, by various adaptive algorithms, the deviations in performance from optimality over time. All prior methods suffer from one or several of the following limitations: restrictive and unverifiable assumptions (e.g., regarding the uncertainties in vehicle dynamics see [4]), limited applicability, and computationally intractable subroutines. In this project, *the first polynomial-time algorithm was developed for the adaptive Linear Quadratic Regulator (LQR) problem that provides high probability guarantees on deviations from optimality (the so-called sublinear regret) and that does not require unverifiable or unrealistic assumptions [4].* This robust adaptive control algorithm was found to:

1. guarantee stability and near-optimal performance at all times;
2. achieve high quality, and
3. produce finite-dimensional, semi-definite programs of size logarithmic in the time horizon (the so-called time horizon from LQR theory is the period of time over which vehicle dynamics are defined).

These properties were demonstrated for a Laplacian system with marginally unstable linear dynamics that has been previously studied in the literature. Uncertainty in the operational environment was introduced using a stable transition matrix; see [4].

Moving from adaptive control to safe execution, the research team studied how to guarantee safe learning in control systems. Data-driven design has considerable potential in contemporary control systems in which precise modeling of the dynamics is intractable, whether due to complex large-scale interactions or nonlinearities resulting from contact forces. However, a large hurdle in the way of practical deployment is the question of maintaining safe operating conditions during the learning process, and guaranteeing safe execution using learned components. Motivated by this issue, the data-driven design of a controller for constrained optimal control was studied [5]. The team designed a controller for a potentially unknown linear dynamical system that minimizes a given cost, subject to the additional requirement that both the state and input stay within a specified *safe region*.

The team directly addressed the tension between exploration for learning and safety, which are fundamentally at odds, by synthesizing a controller that simultaneously excites and regulates the system. System learning was achieved by additively injecting bounded noise to the control inputs computed by a safe controller. By leveraging the recently developed SLS framework for control design, a computationally tractable algorithm was derived [5] that returns a controller, which:

1. guarantees the closed loop system remains within the specified constraint set and
2. ensures that enough noise can be injected into the system to obtain a statistical guarantee on learning.

To the best of our knowledge*, our algorithm is the first to simultaneously achieve both objectives* [5]. This was demonstrated on a small, double integrator example; see [5].

As with all other parts of this project, the controller synthesis was achieved by solving a convex optimization problem, whose solution guarantees the satisfaction of the specified safety constraints on system states and control inputs. The optimization problem was created by lifting the initially non-convex problem into a new parameter representation using SLS, which was convex and tractable to solve while of a high-dimensional character. Also, the reparameterization had the attractive side effect of allowing the performance degradation incurred by meeting the state and input constraints to be understood. This in turn yielded *the first end-to-end sample complexity guarantee for the control of constrained systems* [5].

Finally, the team examined how to use our robust learning framework to incorporate complex, data-driven perception in agile feedback loops. Motivated by vision-based control of autonomous vehicles, the team specifically considered the problem of controlling a known linear dynamical system for which partial state information (e.g., vehicle position) can only be extracted from high-dimensional data (e.g., an image). The approach taken was to learn a perception map from high-dimensional data to partial-state observation and its corresponding error profile, and then to use the perception map in the design of a robust controller. Under suitable smoothness assumptions on the perception map and generative model relating state to high-dimensional data, it was shown [6] an affine error model is sufficiently rich to capture all possible error profiles and can further be learned via a robust regression problem. It was then demonstrated how to integrate the learned perception map and error model into a novel robust control synthesis procedure, and it was further proven that the resulting perception and control loop had favorable generalization properties [6].

The usefulness of the overparameterized approach developed in this project was illustrated in both simulation and a physical deployment. In the synthetic example, we showed how SLS-based control could improve the performance of self-driving in the simulation platform CARLA [6]. This validation step was made before proceeding to physical testing. Figure 1 (left) shows an image generated by CARLA from the virtual perspective of a simulated car. This two-dimensional synthetic data is characteristic of a slice of streaming video data collected by autonomous cars. Figure 1 (right) depicts image analysis of the CARLA image slice that picks out key image features (green boxes) and uses those features for estimating vehicle position (partial state information) using ORB-SLAM2. Simulated use of SLS-based control eliminated vehicle-building contact and met safety (bounds on states) constraints.

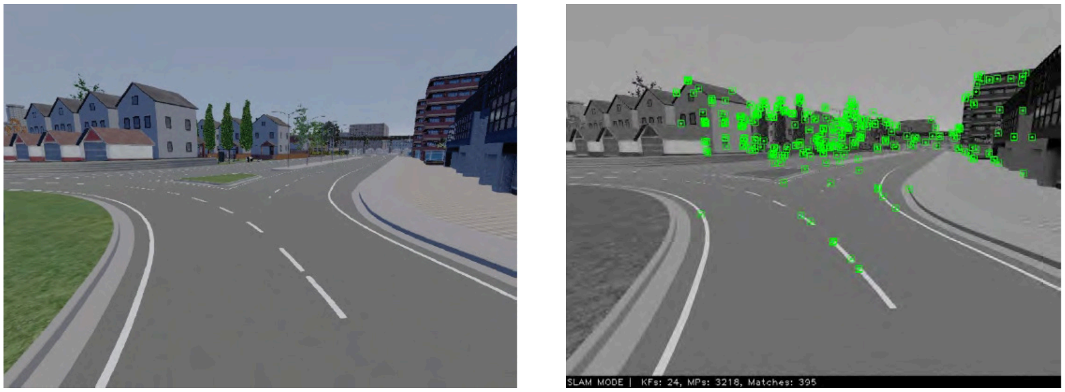

Figure 1. Sample CARLA image (left); features extracted from image data (right).

On a real 1/10th scale car, the SLS framework was implemented to teach a car to drive around an unspecified track from a single lap demonstration performed by a remote driver. See Figure 2. The goal of the trial was for the car to drive as fast as possible around the track specified by the initial lap demonstration, while adapting the control strategy and avoid collision with objects in the environment. As in the CARLA simulation, the challenges were the absence of depth information and the relative nature of estimated position. The car improved its performance over a series of autonomously driven laps while guaranteeing safety (collision avoidance), ending up with faster laps than performed by human trainers.

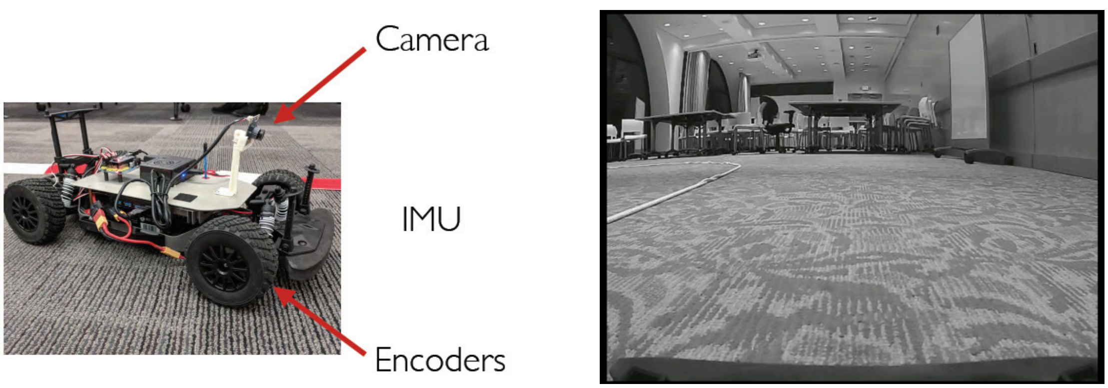

Figure 2. 1/10th scale car with single camera used to collect images real time (left); sample image showing laboratory and rope marking perimeter of track (right).

4. UNDERSTANDING PARAMETERIZATION TRADEOFFS IN RECURRENT NEURAL NETWORKS

The over-arching goal of this research thrust was to reparameterize neural networks to simplify learning. A long-term priority of this research team has been to determine when the output of optimization problems can be well-approximated by recurrent and residual network structures deployed in the practice of neural networks. Here, the question of whether simpler structures may be appropriate to model these maps was investigated, with a focus on understanding parameterization as a resource for problems involving time-series data. In particular, the team developed understanding of how stability could be used as a resource to enable optimization problems that are simple to solve, but powerful enough to be used on state-of-the-art machine learning tasks.

Stability, though a fundamental property of dynamical systems, had, to this date, little bearing on the practice of recurrent neural networks. In work performed in this project, a thorough investigation of stable recurrent models in machine learning was conducted. Recurrent models use feedback to propagate information over time, which allows compact modeling of long-range dependencies in tasks concerning time series and text. However, recurrent models are difficult to train and often exhibit unpredictable behavior. The team investigated whether these recurrent models could be replaced by dynamically stable variants that are easier to train and have more reliable and repeatable behavior at inference time.

Stable recurrent models were studied on a variety established learning benchmarks [9]: language modeling for predicting the next character or word in a sequence (Penn Treebank and Wikitext-2 benchmark datasets); polyphonic music modeling for predicting the next fragment of music in a sequence (JSB Chorales polyphonic benchmark dataset), and slot-filling for interpreting command inputs from word sequences (Airling Travel Information Systems benchmark dataset). Theoretically, it was proved that stable recurrent neural networks are well approximated by feed-forward networks for the purpose of both inference and training by gradient descent [9]. Empirically, it was shown that *stable recurrent models often perform as well as their unstable counterparts on benchmark sequence tasks* [9]. Taken together, these findings shed light on the effective power of recurrent networks and suggest much of sequence learning happens, or can be made to happen, in the stable regime. Moreover, our results help to explain why in many cases practitioners succeed in replacing recurrent models by feed-forward models.

5. GUARANTEED FAST EXECUTION OF FAST NONCONVEX OPTIMIZATION SOLVERS VIA REPARAMETERIZATION

Semi-definite programming (SDP) with equality constraints arise in many optimization and machine learning problems, such as Max-Cut, community detection, and robust PCA. In all of these cases, the SDP formulations are a high dimensional parameterization of a lower dimensional, non-convex problem.

Although SDPs can be solved to arbitrary precision in polynomial time, generic convex solvers do not scale well with the dimension of the problem. In order to address this issue, Burer and Monteiro [2] proposed to reduce the dimension of the problem by appealing to a low-rank factorization, and solve the subsequent non-convex problem instead. It is well-understood that the resulting non-convex problem acts as a reliable surrogate to the original SDP, and can be efficiently solved using the block-coordinate maximization method. Despite its simplicity, remarkable success, and wide use in practice, the theoretical understanding of the convergence of this method is limited.

It was proved in this project [7] that the block-coordinate maximization algorithm applied to the non-convex Burer-Monteiro formulation approach has a global sublinear rate without any assumptions on the problem. It was also proved [7] that this algorithm converges linearly around a local maximum provided that the objective function exhibits quadratic decay. This condition generically holds when the rank of the factorization is sufficiently large. Numerical results back up these theoretical findings.

In this project, progress was also made on developing general purpose methods to avoid *stationary points* in non-convex optimization. Stationary points are neither local minima nor global minima: they are simply places where optimization methods "get stuck" because they cannot find configurations that improve their current optimization cost. We established *that first-order methods (i.e., methods that only use gradient information) avoid saddle points for almost all initializations*. Our results apply to a wide variety of first-order methods, including gradient descent, block coordinate descent, mirror descent and variants thereof. The connecting thread is that such algorithms can be studied from *a dynamical systems perspective* in which appropriate instantiations of the Stable Manifold Theorem allow for a global stability analysis. Thus, neither access to second-order derivative information nor randomness beyond initialization is necessary to provably avoid saddle points.

Even though these methods provably avoid saddle points under random initialization, saddle points can still in practice slow down optimization methods. To fix this problem, a generic framework was developed that generates a sequence of iterates converging to an approximate local minimizer for constrained non-convex problems [10]. The proposed framework consists of two main stages: in the first stage, first-order information is used to reach a first-order stationary point, and in the second stage, second-order information is incorporated to escape from a stationary point if it is a local maximizer or a strict saddle point. These escape directions are efficiently found using sampling tools from machine learning.

Additional research was carried out to solve a class of constrained non-convex, non-concave saddle point problems in a decentralized manner by a group of nodes in a network [8]. With regards to attaining the first-order solution, the algorithm resulting from this research obtained an asymptotic convergence rate proportional to the inverse-square root of N, where N is the number of iterates [8]. This is the first known convergence guarantee for decentralized solution of the saddle point problem and was verified numerically on a general adversarial network.

Additionally, new insights into understanding acceleration of optimization when there are adversarial disturbances were gained in this project [1]. In a meta-turn, *control theory was used to gain an improved understanding of optimization.* Specifically, the trade-offs between convergence rate and robustness to gradient errors in designing a first-order algorithm were studied. Focus was given to gradient descent and accelerated gradient (AG) methods for minimizing strongly convex functions when the gradient has random errors in the form of additive white noise. With gradient errors, the function values of the iterates need not converge to the optimal value; hence, the robustness of an algorithm to noise was defined as the expected suboptimality due to the input noise. For this robustness measure, exact expressions for the quadratic case were derived using tools from robust control theory and tight upper bounds for the smooth strongly convex case using Lyapunov functions certified through matrix inequalities [1]. These characterizations were used within an optimization problem that selects parameters of Multistage Accelerated Stochastic Gradient (M-ASG) algorithms to achieve a particular trade-off between rate and robustness.

Computed results show that M-ASG methods can achieve acceleration while being more robust to random gradient errors; behavior of these practical methods is markedly superior to standard gradient descent and accelerated gradient methods [1]. This was demonstrated for two sets of numerical experiments: the first involving a strongly convex quadratic function with gaussian noise on the function gradients, and the second involving a logistic regression problem of digit image classification with training noise. Some theoretical connections were also established between the robustness of an algorithm and how quickly it can converge back to the optimal solution if it is perturbed from the optimal point with deterministic noise.

6. CONCLUSIONS AND RECOMMENDATIONS

The goal of this fundamental research was to increase the reliability and efficiency of non-convex optimization for large-scale problems representative of challenging environments and uncertain models, and to define the controlled conditions under which this improvement could be achieved. The application targeted by this team was the control of autonomous systems as enabled by machine learning. The intersection of control theory and machine learning has become key to the development of future autonomous systems, as environments, models, and sensing datasets have all increased in complexity beyond the bounds of applicability of existing control methods.

Reparameterization was the primary tool used towards meeting the project goal; this approach yielded faster and quantifiably reliable solution algorithms, with the tradeoff of solving problems in higher dimensional spaces. A framework based on this approach was developed and successfully tested for machine learning of a control system for an autonomous car, which learned how to navigate a track using a single human-operated training lap and partial state, two autonomously operated slow laps to map out safe states, and positional data inferred from an on-board camera. Real-time learning during subsequent laps using the adaptive control framework ultimately yielded much faster laps than performed by human trainers.

Theoretical results were established and the scope of the research broadened beyond vehicle control, as documented in over ten scholarly papers cited in this report. Guarantees were established for the SLS parameterization framework meeting safety constraints on states and inputs subject to environmental uncertainties and computational tractability. The first polynomial-time algorithm of practical utility was developed for the adaptive LQR problem, as well as the first end-to-end sample complexity guarantee for the control of constrained systems. The team also investigated whether recurrent models for tasks concerning time series and streaming text could be replaced by dynamically stable variants that are easier to train and have more reliable and repeatable behavior at inference time. Finally, a variety of fast algorithms were developed for solving non-convex and saddle point problems using reparameterization.

Future work should target the adaptive control of nonlinear, time-varying systems. The models considered herein were assumed linear, a good assumption for many systems. However, for some systems, certain important dynamics may be fundamentally nonlinear and require new control strategies. Research in the community regarding the extension to nonlinear models is ongoing, but has not simultaneously addressed the satisfaction of safety constraints and the efficient use of data samples required for the learning process.

7. REFERENCES

1. N. S. Aybat, A. Fallah, M. Gurbuzbalaban, and A. Ozdaglar. “A Universally Optimal Multistage Accelerated Stochastic Gradient Method.” 2019. arXiv:1901.08022

2. S. Burer and R. D. C. Monteiro. “A Nonlinear Programming Algorithm for Solving Semidefinite Programs via Low-Rank Factorization,” *Mathematical Programming*, 95(2): 329-357, 2003.

3. S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu. “On the Sample Complexity of the Linear Quadratic Regulator.” *Foundations of Computational Mathematics. 2019.* arXiv:1710.01688

4. S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu. “Regret Bounds for Robust Adaptive Control of the Linear Quadratic Regulator.” In *NeurIPS 2018.* arXiv:1805.09388

5. S. Dean, S. Tu, N. Matni, and B. Recht. “Safely Learning to Control the Constrained Linear Quadratic Regulator.” In *Proceedings of the Americal Control Conference 2019.* arXiv:1809.10121

6. S. Dean, N. Matni, B. Recht, and V. Ye. “Robust Guarantees for Perception-Based Control.” Submitted to *Learning for Dynamics and Control*. 2019. arXiv:1907.03680

7. M. A. Erdogdu, A. Ozdaglar, P. A. Parrilo, and N. D. Vanli. “Convergence Rate of Block-Coordinate Maximization Burer-Monteiro Method for Solving Large SDPs.” arXiv:1807.04428

8. W. Liu, A. Mokhtari, A. Ozdaglar, S. Pattathil, Z. Shen, N. Zheng. “A Decentralized Proximal Point-type Method for Saddle Point Problems.” 2019. arXiv:1910.14380

9. J. Miller and M. Hardt. “Stable Recurrent Models.” In *International Conference on Learning Representations.* 2019. openreview.net/forum?id=Hygxb2CqKm and arXiv:1805.10369

10. A. Mokhtari, A. Ozdaglar, A. Jadbabaie. “Escaping Saddle Points in Constrained Optimization.” In *NeurIPS* 2018. arXiv:1809.02162

11. M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, B. Recht. “Learning without Mixing: Towards a Sharp Analysis of Lienar System Identification.” In *Conference on Learning Theory (COLT)* 2018. arXiv:1802.08334

12. Y.-S. Wang, N. Matni, J. C. Doyle. “A System Level Approach to Controller Synthesis.” 2016. arXiv:1610.04815