



**AFRL-AFOSR-VA-TR-2023-0241**

---

**Distributed Coordination in Multi-Agent Networked Systems: Algorithms and Fundamental Limits**

**Na Li  
HARVARD COLLEGE PRESIDENT & FELLOWS OF  
1350 MASS AVE STE 600  
CAMBRIDGE, MA,  
US**

---

**01/04/2023  
Final Technical Report**

**DISTRIBUTION A: Distribution approved for public release.**

Air Force Research Laboratory  
Air Force Office of Scientific Research  
Arlington, Virginia 22203  
Air Force Materiel Command

## REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

<b>1. REPORT DATE</b> 20230104	<b>2. REPORT TYPE</b> Final	<b>3. DATES COVERED</b>	
		<b>START DATE</b> 20171215	<b>END DATE</b> 20211214
<b>4. TITLE AND SUBTITLE</b> Distributed Coordination in Multi-Agent Networked Systems: Algorithms and Fundamental Limits			
<b>5a. CONTRACT NUMBER</b>	<b>5b. GRANT NUMBER</b> FA9550-18-1-0150	<b>5c. PROGRAM ELEMENT NUMBER</b> 61102F	
<b>5d. PROJECT NUMBER</b>	<b>5e. TASK NUMBER</b>	<b>5f. WORK UNIT NUMBER</b>	
<b>6. AUTHOR(S)</b> Na Li			
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> HARVARD COLLEGE PRESIDENT & FELLOWS OF 1350 MASS AVE STE 600 CAMBRIDGE, MA US			<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203		<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> AFRL/AFOSR RTA2	<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b> AFRL-AFOSR-VA-TR-2023-0241
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> A Distribution Unlimited: PB Public Release			
<b>13. SUPPLEMENTARY NOTES</b>			
<b>14. ABSTRACT</b> Multi-agent networked systems arise in many application domains that are of extreme importance to the Air Force. Regardless of the specific application domain, the central goal is to achieve desired collective behavior of the networked system through the design of admissible local agent control algorithms. Despite the large number of algorithms that have been developed for various applications, we lack the knowledge and methods required to deliberately and systematically shape the network collective behavior due to various challenges such as locality, limited communication, and non-stationary time-varying environments. The overarching goal of this project is to formally advance the systematic design of distributed coordination algorithms to ensure high-performance and highconfidence operation of the multiagent networked systems and to investigate fundamental performance limits placed by the various constraints in multi-agent coordination. The project entails three complementary thrusts: 1) developing distributed optimization and control algorithms with rigorous performance guarantees; 2) developing communication efficient schemes; 3) developing online control algorithms to handle both system intrinsic internal uncertainties and external uncertainties			
<b>15. SUBJECT TERMS</b>			
<b>16. SECURITY CLASSIFICATION OF:</b>		<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U	UU 9
<b>19a. NAME OF RESPONSIBLE PERSON</b> FREDERICK LEVE			<b>19b. PHONE NUMBER (Include area code)</b> 696-9730



## **AFOSR Deliverable Reporting Format: Final Reports**

Project: FA9550-18-1-0150  
Distributed Coordination in Multi-Agent Networked Systems: Algorithms  
and Fundamental Limits

**PI: Na Li (Harvard University)**

## Section 1: Structured Survey Questions

### Award Information

- Award Number: FA9550-18-1-0150
- Report Type: Final performance
- Principal Investigator: Na Li
- Principal Investigator Email: nali@seas.harvard.edu
- Principal Investigator Phone: 6263196658
- Project Title: Distributed Coordination in Multi-Agent Networked Systems: Algorithms and Fundamental Limits
- Recipient Organization: HARVARD COLLEGE PRESIDENT & FELLOWS OF
- Business Office Email: awardsmgmt3@harvard.edu
- Report Due Date: 3/14/2022
- Report Period Start Date: 12/15/2020
- Report Period End Date: 12/14/2021
- Current Program Officer: Leve, Frederick

### Publications

Publications during the reporting period: 12/15/2020-12/14/2021

- Yang Zheng, Luca Furieri, Maryam Kamgarpour, Na (Lina) Li, "System-level, Input-output and New Parameterizations of Stabilizing Controllers, and Their Numerical Computation", Automatica, Accepted
- Guannan Qu, Adam Wierman, Na Li, "Scalable Reinforcement Learning for Multi-Agent Networked Systems", Operation Research, ISSN 1526-5463 (online), Accepted
- Yingying Li\*, Yujie Tang\*, Runyu Zhang, Na Li, "Distributed Reinforcement Learning for Decentralized Linear Quadratic Control: A Derivative-Free Policy Optimization Approach", DOI: 10.1109/TAC.2021.3128592, 2021
- Yang Zheng, Na Li, "Non-Asymptotic Identification of Linear Dynamical Systems Using Multiple Trajectories", IEEE Control System Letters, 5 (5), pages 1693-1698, 2021
- Yingying Li, Guannan Qu, Na Li, "Online Optimization with Predictions and Switching Costs: Fast Algorithms and Fundamental Limits", IEEE Transactions of Automatic Control, 66(10): 4761 - 4768, 2021
- Jorge I. Poveda, Na Li, "Robust Hybrid Zero-Order Optimization Algorithms with Acceleration via Averaging in Time", Automatica, volume 123, 109361, 2021.
- Yujie Tang, Junshan Zhang, Na Li, "Distributed Zero-Order Algorithms for Nonconvex Multi-Agent Optimization", IEEE Transactions of Control of Network Systems, 8(1), 269 - 281, 2021.

- Tianpeng Zhang, Victor Qin, Yujie Tang, Na Li, “Source Seeking by Dynamic Source Location Estimation”, IROS, 2021
- Yang Zheng, Yujie Tang, Na Li, “Analysis of the Optimization Landscape of Linear Quadratic Gaussian (LQG) Control”, Learning for Dynamics and Control, 2021
- Yang Zheng, Luca Furieri, Maryam Kamgarpour, Na Li, “Sample Complexity of Linear Quadratic Gaussian (LQG) Control for Output Feedback Systems”, Learning for Dynamics and Control, 2021
- Jiaying Zhou, Xun Xian, Na Li, Jie Ding, “Assisted Learning: Cooperative AI with Autonomy”, ICASSP 2021
- Runyu Zhang, Yingying Li, Na Li, “On the Regret Analysis of Online LQR Control with Predictions”, ACC 2021
- Zhaolin Ren, Aoxiao Zhong, Na Li, “LQR with Tracking: A Zeroth-order Approach and Its Global Convergence”, ACC 2021
- Junshang Zhang, Na Li, Mehmet Dedeoglu, “Federated Learning over Wireless Networks: A Band-limited Coordinated Descent Approach”, IEEE INFOCOM 2021.
- Yingying Li, Subhro Das, Na Li, “Online Optimal Control with Affine Constraints”, AAAI-21, 2021.
- Guannan Qu, Yiheng Lin, Adam Wierman, Na Li, “Scalable Multi-Agent Reinforcement Learning for Networked Systems with Average Reward”, NeurIPS 2020.

## Participants

- Na Li, PI, leads the project and advises students. 1 month
- Yingying Li, Ph.D. student, work on the research of safe online adaptive control and multiagent distributed LQR, 4 months
- Yujie Tang, Postdoc, work on the multiagent zeroth-order feedback optimization and distributed LQR, 4 months
- Runyu Zhang, Ph.D. student, work on the multiagent dynamical systems with heterogeneous interests, 3 months
- Tianpeng Zhang, Ph.D. student, work on building a multi-robot physical lab and a project of multi-robot source seeking, 2 months.

(Estimated number of months\*)

## Other Partners or Collaborators

University of Illinois Urbana-Champaign: Jeff Shamma, Tamer Basar, Bin Hu

California Institute of Technology: Adam Wierman

University of California, San Diego: Yang Zheng

Google Brain: Bo Dai, Jincheng Mei, Schuurmans

IBM: Subhro Das,

### **Awards and Honors (since 2021)**

- 2023 IFAC Thoma Manfred Medal (The Manfred Thoma medal recognizes outstanding contributions of a young researcher and/or engineer under the age of 40 to the field of systems and control in its widest sense. The medal is awarded by the IFAC Council on the recommendation of a selection committee. Citation: For fundamental contributions to control of multiagent networked systems and application to biomedical and energy systems.
- 2020-2023 IFAC Pavel J. Nowacki Distinguished Lecturer.
- 2022 Semi-plenary speaker at International Symposium on Mathematical Theory of Networks and Systems (MTNS2022)
- 2022 Plenary speaker at MIT LIDS Student Conference 2022
- 2022 Plenary speaker at the inaugural workshop of AI Institute on Dynamical Systems
- 2021 Outstanding student paper award at IEEE Conference on Decision and Control (as an adviser)
- 2021- 2026 NSF AI Institute on Dynamical Systems (Thrust leader on Control)
- 2021 Keynote speaker at UIUC CSL Student Conference 2021

### **Former group members status**

1. Former Ph.D. student Guannan Qu who contributed to this project is a tenure-track assistant professor at Carnegie Mellon University.
2. Former postdoc Yang Zheng who contributed to this project starts is a tenure track assistant professor at University of California, San Diego.
3. Former postdoc Sindri Magnusson who contributed to this project is an associate professor in Stockholm University.
4. Former Postdoc Yujie tang who contributed to this project is an assistant professor at Peking University
5. Ph.D. student Yingying Li graduated in July 2021 and will start her tenure track assistant professor appointment at UIUC in Fall 2023.
6. Former undergraduate student Victor Qin who participated in this project is now a Ph.D. student at MIT
7. Former undergraduate student Darrell Huang who participated in this project works at SpaceX starlink.

## Section 2: Technical Report PDF Upload

FA9550-18-1-0150

Distributed Coordination in Multi-agent Networked Systems: Algorithms and Fundamental Limits  
PI: Na Li

Abstract:

Multi-agent networked systems arise in many application domains that are of extreme importance to the Air Force. Regardless of the specific application domain, the central goal is to achieve desired collective behavior of the networked system through the design of admissible local agent control algorithms. Despite the large number of algorithms that have been developed for various applications, we lack the knowledge and methods required to deliberately and systematically shape the network collective behavior due to various challenges such as locality, limited communication, and non-stationary time varying environments. The overarching goal of this project is to formally advance the systematic design of distributed coordination algorithms to ensure high-performance and high-confidence operation of the multiagent networked systems and to investigate fundamental performance limits placed by the various constraints in multi-agent coordination. The project entails three complementary thrusts: 1) developing distributed optimization and control algorithms with rigorous performance guarantees; 2) developing communication efficient schemes; 3) developing online control algorithms to handle both system intrinsic internal uncertainties and external uncertainties

### Activities and Accomplishments, 12/15/2020-12/14/2021

Centered around the three main thrusts, we have made the following progresses:

- i) Multi-agent Distributed Zero-order Nonconvex Feedback Optimization: We continued working on the algorithm design for multiagent feedback optimization. Specifically, we consider a cooperative multi-agent system comprising a group of  $n$  decision-making agents connected by a communication network. Associated with each agent is a local action  $x_i \in R^{d_i}$ , and after the agents take their actions, a local cost  $f(x_1, \dots, x_n)$  will be observed by agent  $i$  which reflects the impact of all agents' actions. The goal for the agents is to cooperatively seek their local actions that minimize their averaged cost as the global objective characterizing the system-wise performance. We study the model-free setting, where each agent can only utilize the observed (zeroth-order) feedback values of the associated local cost, but not (higher-order) derivatives thereof. We propose a Zeroth-order Feedback Optimization (ZFO) algorithm based on local computation and communication of the two-point zeroth-order gradient estimators. More specifically, for each iteration, each agent first takes its own actions and observes the corresponding zeroth-order values of its own local cost, then collects and updates zeroth-order information of other agents' costs by exchanging data with its neighbors in the network, and finally constructs a two-point zeroth-order partial gradient estimate for updating its own action vector. The communication network could be subject to potential delays. We conduct complexity analysis of our ZFO algorithm for both convex and nonconvex objectives, and derive complexity bounds for both noise-free and noisy zeroth-order evaluations. A summary of the complexity bounds can be found in Table 1. Here we list the number of iterations needed for the proposed algorithm to converge with accuracy  $\epsilon > 0$ . These complexity bounds are also compared with the centralized counterparts. In addition, apart from the dependence on  $\epsilon$ ,

we also provide the dependence of the complexity bounds on the problem’s dimension  $d$ , and on the communication network’s structure and delays. To the best of our knowledge, this work is the first to provide explicit complexity bounds for algorithms of multi-agent zeroth-order feedback optimization with analysis on the impact of problem dimension and network structure.

	constrained convex setting		unconstrained nonconvex setting
	$x^* \in \text{int } \mathcal{X}$ known	$x^* \in \text{int } \mathcal{X}$ not known	
noiseless	$\Theta\left(\frac{\bar{b}d}{\epsilon^2}\right)$		$\Theta\left(\frac{\bar{b}\sqrt{nd}}{\epsilon^2}\right)$
noisy	$\Theta\left(\frac{\bar{b}(d^2 + d \ln(1/\epsilon))}{\epsilon^3}\right)$	$\Theta\left(\frac{\bar{b}(d^2 + d \ln(1/\epsilon))}{\epsilon^4}\right)$	$\Theta\left(\frac{\bar{b}\sqrt{nd}^2}{\epsilon^3}\right)$

Table 1: Complexity bounds for our Zeroth-order Feedback Optimization algorithm.

- ii) **Optimal control of partially observable dynamical systems:** One challenge in multiagent control over large-scale systems is that each agent only has a partial observation of the global system. This partial observability places issues for algorithms to converge, not to mention to achieve optimality. Motivated by this, in the past year, we have started investigating the optimization landscape of the optimal control problem with partial observability. Specifically, we revisit the classical Linear Quadratic Gaussian (LQG) control from a modern optimization perspective. We analyze two aspects of the optimization landscape of the LQG problem: 1) connectivity of the set of stabilizing controllers; and 2) structure of stationary points. It is known that similarity transformations do not change the input-output behavior of a dynamic controller or LQG cost. This inherent symmetry by similarity transformations makes the landscape of LQG very rich. We show that 1) the set of stabilizing controllers has at most two path-connected components and they are diffeomorphic under a mapping defined by a similarity transformation; 2) there might exist many strictly suboptimal stationary points of the LQG cost function that are not controllable and not observable; 3) all controllable and observable stationary points are globally optimal and they are identical up to a similarity transformation. These results shed some light on the performance analysis of direct policy gradient methods for solving the LQG problem. We have been working on use these results to derive policy gradient methods that can learn the optimal controller of LQG with a focus on how to escape the non-optimal saddle points.
- iii) **Multiagent reinforcement learning:** In the previous years of efforts, we focused on the cooperative setting for multiagent reinforcement learning where individual agents share a common global objective. Since in practical systems, agents often have different reward functions. In the past year, we have been working on multiagent RL under the game setting where individual agents may share different objectives and aim at designing multiagent RL algorithms that are scalable with respect to the number of agents. Specifically, we study the performance of the gradient play algorithm for stochastic games (SGs), where each agent tries to maximize its own total discounted reward by making decisions independently based on

current state information which is shared between agents. Policies are directly parameterized by the probability of choosing a certain action at a given state. We show that Nash equilibria (NEs) and first-order stationary policies are equivalent in this setting, and give a local convergence rate around strict NEs. Further, for a subclass of SGs called Markov potential games (which includes the cooperative setting with identical rewards among agents as an important special case), we design a sample-based reinforcement learning algorithm and give a non-asymptotic global convergence rate analysis for both exact gradient play and our sample-based learning algorithm. Our result shows that the number of iterations to reach an  $\epsilon$ -NE scales linearly, instead of exponentially, with the number of agents. Local geometry and local stability are also considered, where we prove that strict NEs are local maxima of the total potential function and fully-mixed NEs are saddle points.

- iv) Limited communication in distributed coordination: In the past year, we have been continuing to work on the distributed optimization with limited communication bandwidth. Specifically, we consider large scale distributed optimization over a set of workers connected to a central server where the limited communication bandwidth between the server and workers imposes a significant bottleneck for the optimization procedure. We propose a distributed stochastic gradient descent type algorithm that employs compressed sensing for gradient compression. At the heart of our proposed algorithm, we use compressed sensing techniques for the compression of local stochastic gradients at the worker side and the reconstruction of the aggregated stochastic gradients at the server side. The use of compressed sensing enables the server to approximately identify the top entries of the aggregated gradient without querying directly each local gradient. We theoretically prove that our algorithm achieves comparable convergence with the vanilla SGD even in the presence of additive noise incurred by the uploading communication channels. Our algorithm is inspired by and goes beyond the literature that uses sketching schemes for gradient compression. In addition, our methods could handle transmission error/noise of the uploading links.
- v) Safe online control with time-varying cost functions: Previously, we had been only working on online optimal control without constraints on the state and controller. However, in many real-world systems, both system states and controllers have safety constraints, e.g., should stay in a safe set always. In this report period, we have designed safe online control algorithms with time-varying cost functions and safe constraints. Specifically, we design safe algorithms for time-varying environments by considering the following constrained online optimal control problem  $x_{t+1} = f(x_t, u_t, w_t)$  where  $x_t$  is the state,  $u_t$  is the control action, and  $w_t$  is bounded disturbances. We consider affine constraints  $D_t x_t \leq d_x$  and  $D_u u_t \leq d_u$ . At each stage, there is also a cost function  $c_t(x_t, u_t)$  is adversarially generated and the decision maker selects a feasible action  $u_t$  before  $c_t(x_t, u_t)$  is revealed. We aim to achieve two goals simultaneously, i) to minimize the sum of the adversarially varying costs with sublinear regrets, ii) to satisfy the constraints for all  $t$  despite the disturbances. To solve this problem, we propose Online Gradient Descent with Buffer Zones (OGD-BZ). Theoretically, we show that OGD-BZ with proper parameters can guarantee the system to satisfy all the constraints despite any admissible disturbances. Further, we investigate the policy regret of OGD-BZ, which compares OGD-BZ's performance with the performance of the optimal linear policy in hindsight. We show that OGD-BZ can achieve a policy regret upper bound that is square root of the horizon length multiplied by some logarithmic terms of the horizon length under proper algorithm

parameters. As far as we know, OGD-BZ is the first algorithm with theoretical guarantees on both sublinear policy regret and robust constraint satisfaction.

## Impacts

### **Development of the principal discipline(s) of the project**

The major results have been illustrated in the previous section. Here we only provide a brief summary for each line of work with emphasis on the significance of the results to applications.

- i) Our safe online optimal control algorithm is among the first that can learn unknown models, satisfy safety constraints despite random disturbances and achieves sublinear policy regret with rigorous provable guarantee, indicating a good online performance.
- ii) Our multiagent feedback zero-order nonconvex optimization algorithms introduces a new multiagent optimization problem framework where agents can only evaluate the outcomes of joint actions by taking physical actions. It has many practical DOD applications especially robotic swarms.
- iii) Our recent limited communication in distributed coordination reduce large dimension of communication needs to low dimension through compressive sensing. This tool can be combined with the quantization scheme we developed in the early years of this project for realizing a practical communication scheme which can work under various communication environment.
- iv) Our scalable reinforcement learning methods are among the first scalable reinforcement learning methods for large scale multiagent network systems with rigorous theoretical guarantees in both sample complexity and stability. It opens up broad applications for large scale complex dynamical systems.
- v) Our work on characterizing the optimization landscape of the partially observable systems provide foundations for developing and analyzing policy optimization methods for the control of partially observable systems, one of the most daunting challenges in the control of real-world dynamical systems.
- vi) Our new work on multi-robot learning and control has directly applications to many DOD areas including source seeking, rescuing, surveillance, etc.

### **Other disciplines:**

This project itself is very interdisciplinary, drawing tools from and making progress to the areas of dynamics and control, optimization, reinforcement learning, and robotics. The rich set of algorithms developed in this project has a broad application to different areas, including multi-robots, transportation, energy systems, and neuron network.

### **Describe the impact in this reporting period on the development of human resources**

The project has attracted many strong students, both graduate and undergraduate students, to pursue the interdisciplinary area of control, dynamical systems, optimization, reinforcement learning, and robotics. Below is a list of where the former group members are to show the impact of this project.

1. Former Ph.D. student Guannan Qu who contributed to this project is a tenure-track assistant professor at Carnegie Mellon University.
2. Former postdoc Yang Zheng who contributed to this project starts is a tenure track assistant professor at University of California, San Diego.
3. Former postdoc Sindri Magnusson who contributed to this project is an associate professor in Stockholm University.
4. Former Postdoc Yujie tang who contributed to this project is an assistant professor at Peking University
5. Ph.D. student Yingying Li graduated in July 2021 and will start her tenure track assistant professor appointment at UIUC in Fall 2023.
6. Former undergraduate student Victor Qin who participated in this project is now a Ph.D. student at MIT
7. Former undergraduate student Darrell Huang who participated in this project works at SpaceX starlink.

**Describe the impact on teaching and educational experiences**

The PI has designed and offered a new graduate student control course which draws research results from this project and trains the next generation of researchers in the disciplinary area of area of control, dynamical systems, optimization, reinforcement learning, and robotics

**Describe the impact in this reporting period on physical, institutional, and information resources that form infrastructure.**

Through this project, we have developed robotics lab with a motion capture system using Optitrack cameras and two types of vehicles, two-wheel-ground vehicles, Turtlebots, and small quadrotors, crazyflies. The physical lab not only provides a testbed for the PI's research group to close the loop between theory and applications, but also draw a growing interest in the undergraduate students to pursue the career in control and robotics.