REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
21-06-2018	Final Report	1-Jun-2016 - 28-Feb-2017

4. TITLE AND SUBTITLE
Final Report: Tensor methods for large-scale learning

5a. CONTRACT NUMBER
W911NF-16-1-0134

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
University of California - Irvine
141 Innovation Drive, Suite 250

Irvine, CA 92697 -7600

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
68283-NS-II.5

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Animashree Anandkumar
UU	UU	UU	UU		19b. TELEPHONE NUMBER 949-824-9072

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

RPPR Final Report
as of 25-Jun-2018

Agency Code:

Proposal Number: 68283NSII **Agreement Number: W911NF-16-1-0134**

INVESTIGATOR(S):

Name: Animashree Anandkumar
Email: a.anandkumar@uci.edu
Phone Number: 9498249072
Principal: Y

Organization: **University of California - Irvine**
Address: 141 Innovation Drive, Suite 250, Irvine, CA 926977600
Country: USA

DUNS Number: 046705849 EIN: 952226406
Report Date: 28-Mar-2017 Date Received: 21-Jun-2018
Final Report for Period Beginning 01-Jun-2016 and Ending 28-Feb-2017
Title: Tensor methods for large-scale learning
Begin Performance Period: 01-Jun-2016 **End Performance Period:** 28-Feb-2017
Report Term: 0-Other
Submitted By: Animashree Anandkumar Email: a.anandkumar@uci.edu
Phone: (949) 824-9072
Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: 3 **STEM Participants:**

Major Goals: To explore tensor methods for learning
To develop scalable algorithms that make use of tensor algebra
To analyze performance of algorithms in theory and in practice

Accomplishments: In the pdf document

Training Opportunities: Graduate student Kamyar Azzizade conducted research on reinforcement learning with spectral methods under my guidance as part of this project

Results Dissemination: This was disseminated through workshop and talks

Honors and Awards: The PI was recently selected as a named professor at Caltech.

Protocol Activity Status:

Technology Transfer: We have released many open source software packages based on the tensor algorithms developed as part of this project. We have contributed to Tensorly package \url{https://github.com/tensorly} and actively developing it. The tensor decomposition algorithm for topic modeling is now deployed at AWS SageMaker platform and is part of topic modeling framework in AWS Comprehend, the NLP service.

PARTICIPANTS:

Participant Type: PD/PI
Participant: Animashree Anandkumar
Person Months Worked: 9.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: RLDM 2017
Date Received: Conference Date: 12-Jun-2017 Date Published: 12-Jun-2017
Conference Location: Ann Arbor, Michigan, USA
Paper Title: Reinforcement Learning in Rich-Observation MDPs using Spectral Methods
Authors: Kamyar Azizzadenesheli, Alessandro Lazaric, Anima Anandkumar
Acknowledged Federal Support: **Y**

WEBSITES:

URL: http://tensorly.org/stable/index.html
Date Received: 21-Jun-2018
Title: Tensorly package
Description: Open source package for tensor operations
URL: http://tensorly.org/stable/index.html
Date Received: 21-Jun-2018
Title: Tensorly package
Description: Open source package for tensor operations

Final Report for ARO STIR W911NF-16-1-0134

Prof. Anima Anandkumar

Motivation

Today we are facing a "data deluge" in almost every domain. Online social networks have seen an explosion in activity and have fundamentally transformed the nature of human interaction. In the biological realm, modern genome sequencers can output data at a rate 400 times faster than the ones a decade ago, and so on. However, although having a transformative potential, the data deluge has not yet been exploited to the fullest extent. Ironically, the data deluge has also resulted in a "data desert". The collected data in many domains are noisy, subsampled, with typically a large number of variables or "unknowns" compared to the number of observations or the "knowns". Such high-dimensionality entails practical principled approaches for learning from ill-posed and ill-behaved data.

Some of the fundamental questions in high-dimensional learning are: Can we design **scalable** models for efficiently representing and learning high-dimensional data? Here, scalability refers to low **computational** requirements and reduced **sampling** of high-dimensional data. Not all phenomena can be learnt in a scalable manner. Can we characterize the **fundamental limits** on complexity of learning complex phenomena? If these processes need to be actively sampled and decisions need to be made, based on those observations, can we design efficient reinforcement learning strategies?

As part of this project, the PI Anandkumar has tackled the above challenges by exploiting "inherent data architecture". This can be in the form of structural relationships among the variables, represented as **graphs**, or as parametric forms, represented as **tensor decompositions**. The PI Anandkumar has developed novel approaches for handling such high-dimensional data. Below, the technical approach and outcome for various research topics under the ARO STIR project are described.

1 Spectral methods for reinforcement learning

Starting point: Unsupervised Learning of Latent Variable Models. Learning general latent variable models through maximum likelihood is NP-hard. Previous methods with theoretical consistency guarantees have high computational and sample complexity which typically scale exponentially with the latent space dimensionality. The current practice for estimating latent variable models is mostly through local search heuristics (e.g., the EM algorithm) which are prone to failure in high dimensions.

Classically, latent variables have been incorporated via mixture models. A mixture model can be thought of as selecting the distribution of the observed variables, based on a so-called latent choice variable. **Gaussian mixtures** are the most well studied class of mixture models. Recently the so-called class of **exchangeable topic models** such as **latent Dirichlet allocation** have been popular for modeling large word corpora. These models incorporate documents with multiple hidden topics. The PI proposes efficient methods for learning a wide range of mixture models in [1]. In addition, recently, the PI has considered non-parametric extensions of the tensor approach which enables learning of a wider class of latent variable models [2].

Background on Reinforcement Learning: Reinforcement learning (RL) framework studies the problem of efficient agent-environment interaction, where the agent learns to maximize a given reward function in the long run [3]. At the beginning of the interaction, the agent is uncertain about the environment's dynamics and must *explore* different policies in order to gain information about it. Once the agent is fairly certain, the knowledge about the environment can be *exploited* to compute a good policy attaining a large cumulative reward. Designing algorithms that achieve an effective trade-off between exploration and exploitation is the primary goal of reinforcement learning. The trade-off is commonly measured in terms of *cumulative regret*, that is the difference between the rewards accumulated by the optimal policy (which requires exact knowledge of the environment) and those obtained by the learning algorithm.

In practice, we often deal with environments with large observation state spaces (e.g., robotics). In this case the regret of standard RL algorithms grows quickly with the size of the observation state space. (We use observation state and observation interchangeably.) Nonetheless, in many domains there is an underlying low dimensional latent space that summarizes the large observation space and its dynamics and rewards. For instance, in robot navigation, the high-dimensional visual and sensory input can be summarized into a 2D position map, but this map is typically unknown. This makes the problem challenging, since it is not immediately clear how to exploit the low-dimensional latent structure to achieve low regret.

Summary of contributions: In a recent work [4], as part of the STIR project, we employed rich-observation Markov decision processes (ROMDP), where hidden states are mapped to observations through an injective mapping, so that an observation can be generated by only one hidden state. When this mapping is unknown a priori, we introduce a spectral decomposition method that consistently estimates how observations are clustered in the hidden states. The estimated clustering is then integrated into an optimistic algorithm for RL (UCRL), which operates on the smaller clustered space. The resulting algorithm proceeds through phases and we show that its per-step regret (i.e., the difference in cumulative reward between the algorithm and the optimal policy) decreases as more observations are clustered together and finally matches the (ideal) performance of an RL algorithm running directly on the hidden MDP.

We focus on rich-observation Markov decision processes, where a small number of X hidden states are mapped to a large number of Y observations through an injective mapping, so that an observation can be generated by only one hidden state and hidden states can be viewed as clusters.

In this setting, we show that it is indeed possible to devise an algorithm that starting from observations can progressively cluster them in "smaller" states and eventually converge to the hidden MDP. We introduce algorithm, where we integrate spectral decomposition methods into the upper-bound for RL algorithm (UCRL). The algorithm proceeds in epochs in which an estimated mapping between observations and hidden state is computed and an optimistic policy is computed on the MDP (called auxiliary MDP) constructed from the samples collected so far and the estimated mapping. The mapping is computed using spectral decomposition of the tensor associated to the observation process.

We prove that this method is guaranteed to correctly "cluster" observations together with high probability. As a result, the dimensionality of the auxiliary MDP decreases as more observations are clustered, thus making the algorithm more efficient computationally and more effective in finding good policies. Under transparent and realistic assumptions, we derive a regret bound showing that the per-step regret decreases over epochs, and we prove a worst-case bound on the number of steps (and corresponding regret) before the full mapping between states and observations is computed. The regret accumulated over this period is actually constant as the time to correct clustering does not increase with the number of steps N. As a result, our algorithm asymptotically matches the regret of learning directly on the latent MDP. We also notice that the improvement in the regret comes with an equivalent reduction in time and space complexity. In fact, as more observations are clustered, the space to store the auxiliary MDP decreases and the complexity of the extended value iteration step in UCRL decreases from $O(Y^3)$ down to $O(X^3)$.

Transitions: We have released many open source software packages based on the tensor algorithms developed as part of this project. We have contributed to Tensorly package `https://github.com/tensorly` and actively developing it. The tensor decomposition algorithm for topic modeling is now deployed at AWS SageMaker platform and is part of topic modeling framework in AWS Comprehend, the NLP service.

2 Recent Awards

The PI was recently selected to be a **named professor at Caltech** in 2017, viz., the Bren professorship. It is the Institute's most distinguished award for individual faculty. This honor provides faculty with additional funds and resources to pursue their best ideas while continuing to mentor future generations of leaders. Each named professorship brings its own distinct legacy. Many professorships, for instance, have long-standing histories, and pass on through each appointment a tradition of discovery and exploration from one academic generation to the next, from one colleague to another. A professorship may also provide a faculty member with an opportunity to forge meaningful connections with the philanthropists who provided the donation that made it possible.

Prof. Anandkumar's contributions have been recognized in a number of venues. She was selected for **Microsoft faculty fellowship** (2013). This prestigious award is given to five promising faculty members in the first three years of their faculty career from North America and Europe. She is also the recipient of the **CAREER Award** by the NSF. She also received the Google faculty award.

References Cited

[1] Animashree Anandkumar, Rong Ge, Daniel Hsu, Sham M Kakade, and Matus Telgarsky. Tensor decompositions for learning latent variable models. *The Journal of Machine Learning Research*, 15(1):2773–2832, 2014.

[2] L. Song, A. Anandkumar, B. Dai, and B. Xie. Nonparametric Estimation of Multi-View Latent Variable Models. *ArXiv:1311.3287*, Nov. 2013.

[3] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[4] Kamyar Azizzadenesheli, Alessandro Lazaric, and Animashree Anandkumar. Reinforcement Learning in Rich-Observation MDPs using Spectral Methods. In *RLDM*, 2017.