

# Tolga Ergen

✉ [ergen@stanford.edu](mailto:ergen@stanford.edu) | 🏠 [stanford.edu/~ergen/](https://stanford.edu/~ergen/) | 🐦 [@tolgaergen\\_](https://twitter.com/tolgaergen_)

## Research Interests

Machine learning, deep learning, optimization

## Education

### Stanford University

PH.D. IN ELECTRICAL ENGINEERING, CGPA: **4.11 / 4.00**

- Advisor: Mert Pilanci
- Thesis: Convex optimization for neural networks
- Expected graduation: 06/2023

*Stanford, CA*

*Sep 2018 – Present*

### Bilkent University

M.S. IN ELECTRICAL AND ELECTRONICS ENGINEERING, CGPA: **4.00 / 4.00**

- Advisor: Suleyman Serdar Kozat
- Thesis: Online learning with recurrent neural networks

*Ankara, Turkey*

*Sep 2016 – July 2018*

### Bilkent University

B.S. IN ELECTRICAL AND ELECTRONICS ENGINEERING, CGPA: **3.97 / 4.00**

- Graduated as the **3rd** in class

*Ankara, Turkey*

*Sep 2011 – July 2016*

## Industrial & Academic Experience

### Google Research

RESEARCH INTERN

- Hosts: Harsh Mehta and Behnam Neyshabur
- Improving optimization and understanding of transformer networks through convex optimization theory

*Mountain View, CA*

*Jun 2022 – Sep 2022*

### Salesforce Research

RESEARCH INTERN

- Host: Yu Bai
- Gradient based methods for uncertainty quantification under unknown distribution shift

*Palo Alto, CA*

*Jun 2021 – Sep 2021*

### Stanford University

TEACHING ASSISTANT

- EE-269: Signal Processing for Machine Learning
- EE-270: Large Scale Matrix Computation, Optimization and Learning
- EE-364B: Convex Optimization II
- EE-269: Signal Processing for Machine Learning
- EE-270: Large Scale Matrix Computation, Optimization and Learning
- EE-364B: Convex Optimization II
- EE-364B: Convex Optimization II

*Stanford, CA*

*2018-Present*

*Sep 2019 – Dec 2019*

*Jan 2020 – Mar 2020*

*Apr 2020 – Jun 2020*

*Sep 2020 – Dec 2020*

*Jan 2021 – Mar 2021*

*Mar 2021 – Jun 2021*

*Mar 2022 – Jun 2022*

### Bilkent University

TEACHING ASSISTANT

- EEE-424: Digital Signal Processing
- EEE-102: Introduction to Digital Circuit Design
- EEE-424: Digital Signal Processing
- EEE-424: Digital Signal Processing

*Ankara, Turkey*

*2016-2018*

*Sep 2016 – Jan 2017*

*Feb 2017 – Jun 2017*

*Sep 2017 – Jan 2018*

*Feb 2018 – Jun 2018*

### Academic Service

REVIEWER

- NeurIPS, ICML, ICLR, IEEE Transactions on Neural Networks and Learning Systems (TNNLS) and IEEE Signal Processing Letters (SPL)

### Havelsan Inc., CCCS

UNDERGRADUATE INTERN

- Modelling sound propagation and design of a wireless communication system using Snap modules

*Ankara, Turkey*

*Aug 2015 – Sep 2015*

### Aselsan Inc., REHIS

UNDERGRADUATE INTERN

- Design and implementation of an audio radar warning system based on Virtex-5 FPGA

*Ankara, Turkey*

*Aug 2014 – Sep 2014*

## Honors & Awards

---

2022	Awarded Adobe Research Fellowship	Stanford, CA
2021	Received NeurIPS 2021 Outstanding Reviewer Award given to <b>the top 8% of reviewers</b>	Stanford, CA
2021	International Conference on Acoustics, Speech, & Signal Processing (ICASSP), <b>best paper award</b>	Stanford, CA
2020	Conference on the Mathematical Theory of Deep Neural Networks (DeepMath), <b>best poster award</b>	Stanford, CA
2018	Stanford University Departmental Fellowship: Full tuition waiver & stipend during the first year of PhD	Stanford, CA
2017	Bilkent University Graduate Research Conference (GRC), <b>best oral presentation award</b> in signal processing	Ankara, Turkey
2016	TUBITAK Scholarship for the M.S. studies based on a weighted ALES (National GRE) and GPA score list	Ankara, Turkey
2016	Full Scholarship from Bilkent University during M.S. Studies	Ankara, Turkey
2016	Bilkent University Academic Excellence Award	Ankara, Turkey
2016	Bilkent University High Honor Student during B.S. Studies	Ankara, Turkey
2015	Received the <b>13th rank</b> among 0.2M university graduates in ALES (National GRE)	Ankara, Turkey
2011	Bilkent University Full Scholarship for the B.S. degree in the EEE Department	Ankara, Turkey
2011	Received the <b>178th rank</b> among 2M high school graduates in University Entrance Examinations	Ankara, Turkey

## Publications

---

### JOURNAL ARTICLES

#### Convex Geometry and Duality of Over-parameterized Neural Networks

T. Ergen, M. Pilanci  
*Journal of Machine Learning Research (JMLR)* (2021)

#### A Novel Distributed Anomaly Detection Algorithm based on Support Vector Machines

T. Ergen, S. S. Kozat  
*Elsevier Digital Signal Processing* (2020)

#### Unsupervised Anomaly Detection with LSTM Neural Networks

T. Ergen, S. S. Kozat  
*IEEE Transactions on Neural Networks and Learning Systems* (2019)

#### Energy-Efficient LSTM Networks for Online Learning

T. Ergen, Ali H Mirza, S. S. Kozat  
*IEEE Transactions on Neural Networks and Learning Systems* (2019)

#### Team-optimal Online Estimation of Dynamic Parameters over Distributed Tree Networks

O. F. Kilic, T. Ergen, M. Sayin, S. S. Kozat  
*Elsevier Signal Processing* (2019)

#### Online Training of LSTM Networks in Distributed Systems for Variable Length Data Sequences

T. Ergen, S. S. Kozat  
*IEEE Transactions on Neural Networks and Learning Systems* (2017)

#### Efficient Online Learning Algorithms based on LSTM Neural Networks

T. Ergen, S. S. Kozat  
*IEEE Transactions on Neural Networks and Learning Systems* (2017)

### PREPRINTS

#### Fixing the NTK: From Neural Network Linearizations to Exact Convex Programs

T. Ergen, R. Dwaraknath, M. Pilanci  
*Under Review* (2022)

#### Convexifying Transformers: Improving optimization and understanding of transformer networks

T. Ergen, B. Neyshabur, H. Mehta  
*Under Review* (2022)

#### Scaling Convex Neural Networks with Burer-Monteiro Factorization

A. Sahiner, T. Ergen, B. Ozturkler, J. Pauly, M. Mardani, M. Pilanci  
*Under Review* (2022)

#### Path Regularization: A Convexity and Sparsity Inducing Regularization for Parallel ReLU Networks

T. Ergen, M. Pilanci  
*Under Review* (2021)

## Two-Layer Neural Networks as Sparse Mixtures of Convex Models: Polynomial-Time Convex Optimization Formulations of Neural Networks with Piecewise Linear Activations

T. Ergen, M. Pilanci  
*Under Review* (2021)

## CONFERENCE & WORKSHOP PAPERS

### Globally Optimal Training of Neural Networks with Threshold Activation Functions

T. Ergen, H. Gulluk, J. Lacotte, M. Pilanci  
*International Conference on Learning Representations (ICLR)* (2023)

### Parallel Deep Neural Networks Have Zero Duality Gap

Y. Wang, T. Ergen, M. Pilanci  
*International Conference on Learning Representations (ICLR)* (2023)

### Unraveling Attention via Convex Duality: Analysis and Interpretations of Vision Transformers

A. Sahiner, T. Ergen, B. Ozturk, J. Pauly, M. Mardani, M. Pilanci  
*International Conference on Machine Learning (ICML)* (2022)

### Demystifying Batch Normalization in ReLU Networks: Equivalent Convex Optimization Models and Implicit Regularization

T. Ergen\*, A. Sahiner\*, B. Ozturk, J. Pauly, M. Mardani, M. Pilanci  
*International Conference on Learning Representations (ICLR)* (2022)

### Hidden Convexity of Wasserstein GANs: Interpretable Generative Models with Closed-Form Solutions

A. Sahiner\*, T. Ergen\*, B. Ozturk, B. Bartan, J. Pauly, M. Mardani, M. Pilanci  
*International Conference on Learning Representations (ICLR)* (2022)

### Revealing the Structure of Deep Neural Networks via Convex Duality

T. Ergen, M. Pilanci  
*International Conference on Machine Learning (ICML)* (2021)

### Global Optimality Beyond Two Layers: Training Deep ReLU Networks via Convex Programs

T. Ergen, M. Pilanci  
*International Conference on Machine Learning (ICML)* (2021)

### Implicit Convex Regularizers of CNN Architectures: Convex Optimization of Two- and Three-Layer Networks in Polynomial Time

T. Ergen, M. Pilanci  
*International Conference on Learning Representations (ICLR)*-**(Spotlight Presentation)** (2021)

### Vector-output ReLU Neural Network Problems are Copositive Programs: Convex Analysis of Two Layer Networks and Polynomial-time Algorithms

A. Sahiner, T. Ergen, J. Pauly, M. Pilanci  
*International Conference on Learning Representations (ICLR)* (2021)

### Exact and Relaxed Convex Formulations for Shallow Neural Autoregressive Models

V. Gupta, B. Bartan, T. Ergen, M. Pilanci  
*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*-**(Outstanding Paper Award)** (2021)

### Neural Networks are Convex Regularizers: Exact Polynomial-time Convex Optimization Formulations for Two-layer Networks

M. Pilanci, T. Ergen  
*International Conference on Machine Learning (ICML)* (2020)

### Convex Geometry of Two-Layer ReLU Networks: Implicit Autoencoding and Interpretable Models

T. Ergen, M. Pilanci  
*International Conference on Artificial Intelligence and Statistics (AISTATS)* (2020)

### Convex Programs for Global Optimization of Convolutional Neural Networks in Polynomial-Time

T. Ergen, M. Pilanci  
*NeurIPS Workshop on Optimization for Machine Learning (OPTML)*-**(Oral Presentation)** (2020)

### Random Projections for Learning Non-convex Models

T. Ergen, M. Pilanci  
*NeurIPS Workshop on Beyond First Order Methods in Machine Learning* (2019)

### Convex Duality and Cutting Plane Methods for Over-parameterized Neural Networks

T. Ergen, M. Pilanci  
*NeurIPS Workshop on Optimization for Machine Learning (OPTML)* (2019)

### Convex Optimization for Shallow Neural Networks

T. Ergen, M. Pilanci

Annual Allerton Conference on Communication, Control, and Computing (*Allerton*) (2019)

### A Highly Efficient Recurrent Neural Network Architecture for Data Regression

T. Ergen, E. Ceyani

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2018)

### A Novel Anomaly Detection Approach Based on Neural Networks

T. Ergen, M. Kerpici

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2018)

### Computationally Efficient Online Regression via LSTM Neural Networks

T. Ergen, S. S. Kozat

European Signal Processing Conference (*EUSIPCO*) (2017)

### An Efficient Bandit Algorithm for General Weight Assignments

K. Gokcesu, T. Ergen, S. Ciftci, S. S. Kozat

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2017)

### Neural Networks Based Online Learning

T. Ergen, S. S. Kozat

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2017)

### Novelty Detection Using Soft Partitioning and Hierarchical Models

T. Ergen, K. Gokcesu, M. Simsek, S. S. Kozat

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2017)

### Online Distributed Nonlinear Regression via Neural Networks

T. Ergen, S. S. Kozat

IEEE Signal Processing and Communications Applications Conference (*SIU*) (2017)

## Skills

---

**Programming** Python, Matlab, LaTeX, VHDL

**Languages** English, Turkish