

# Eric Wong

Phone: +1 (339) 223-7159

Email: [wongeric@mit.edu](mailto:wongeric@mit.edu)

Website: <https://riceric22.github.io>

Google Scholar: [\[link\]](#)

## Briefly

---

How can we make sure that deep learning models are actually doing what we want them to do? My research interests are centered around foundations of reliable machine learning systems: understanding, debugging, and guaranteeing the behavior of data-driven models. I created the first provable defenses that guarantee robustness to adversarial examples and real-world specifications.

## Education

---

<b>Massachusetts Institute of Technology, Post-Doctoral Associate</b> <i>Advisor: Aleksander Madry</i>	Cambridge, MA 2020-current
<b>Carnegie Mellon University, Ph. D. in Machine Learning</b> Thesis: Provable, structured, and efficient methods for robustness of deep networks to adversarial examples; SCS Dissertation Award — Honorable Mention <i>Advisor: Zico Kolter</i>	Pittsburgh, PA 2015-2020
<b>Carnegie Mellon University, B. S. in Computer Science</b> Double major in Mathematics, minor in Machine Learning	Pittsburgh, PA 2011-2015

## Selected publications

---

2021	<b>ICML</b>	<b>Leveraging sparse linear layers for debuggable deep networks</b> Eric Wong, Shibani Santurkar, and Aleksander Madry <i>International Conference on Machine Learning</i>
	<b>ICLR</b>	<b>Learning perturbation sets for robust machine learning</b> Eric Wong and J Zico Kolter <i>International Conference on Learning Representations</i>
2020	<b>ICML</b>	<b>Overfitting in adversarially robust deep learning</b> Leslie Rice, Eric Wong, and Zico Kolter <i>International Conference on Machine Learning</i>
	<b>ICLR</b>	<b>Fast is better than free: Revisiting adversarial training</b> Eric Wong, Leslie Rice, and J. Zico Kolter <i>International Conference on Learning Representations</i>
2019	<b>ICML</b>	<b>Wasserstein adversarial examples via projected sinkhorn iterations</b> Eric Wong, Frank Schmidt, and Zico Kolter <i>International Conference on Machine Learning</i>

- 2018 **NeurIPS** **Scaling provable adversarial defenses**  
Eric Wong, Frank Schmidt, Jan Hendrik Metzen, and J. Zico Kolter  
*Advances in Neural Information Processing Systems*
- ICML** **Provable defenses against adversarial examples via the convex outer adversarial polytope**  
Eric Wong and Zico Kolter  
*International Conference on Machine Learning*

## Teaching experience

---

- 2016 **Practical Data Science**  
Designed new assignments, taught recitations, and prepared write-ups for the first iteration of CMU's Practical Data Science course (15-388/688). I was the head TA (out of two TAs) and managed over 300 enrolled students. Received 55 student reviews with an average rating of 4.85/5.
- 2016-2019 **Eberly Center**  
Enrolled in teaching seminars at the Eberly Center in CMU to develop personal teaching skills; seminars include "Teaching Inclusively: Leveraging Diversity and Promoting Equity in Your Classroom" and "Helping Students Develop Mastery and Critical Thinking"
- 2015 **Advanced Introduction to Machine Learning**  
Taught recitations, held office hours, and created/graded assignments for the second iteration of the Advanced Introduction to Machine Learning course intended for doctoral students in CMU's Machine Learning Department.
- 2014 **Algorithm Design and Analysis**  
Taught recitations, conducted oral examinations/office hours, and graded assignments/exams for the computer science department's Algorithm Design and Analysis course (15-451) at CMU.
- 2013, 2014 **Pervasive and Mobile Computing Services**  
Held office hours and graded assignments/exams for the software engineering department's Pervasive and Mobile Computing Services course (08-766/781) at CMU.
- 2013 **Mobile Development for iOS and Android**  
Held office hours and graded assignments/exams for the software engineering department's Mobile Development course (08-723) at CMU.

## Work experience

---

- 2019-2020 **Bosch Center for Artificial Intelligence (Renningen, Germany and Pittsburgh, PA)**  
Created a virtual sensor based on neural networks for a fuel injection system in truck engines; *formally verified the worst-case error of the system* under conservative estimates of physical sensor noise.
- 2012-2015 **CERT Program (Pittsburgh, PA)**  
Migrated secure coding rules from POSIX to C11; analyzed security reports for Java android applications; developed an analysis tool for security vulnerabilities in source code.

## Community service, outreach, and mentorship

---

- 2021 **MIT Undergraduate Research Opportunities Program**  
Directly supervised an undergraduate for the UROP program at MIT; provided an opportunity for a member of an under-represented group to learn about machine learning and tackle a challenging research project  
<https://uop.mit.edu/>
- 2021 **MENTorEd Opportunities in Research (METEOR Program, MIT)**  
Participated in the METEOR postdoc fellowship selection committee, an on-going effort at CSAIL MIT to increase diversity, equity, and inclusion. Provided confidential technical feedback on candidates based on their application materials.  
<https://www.csail.mit.edu/meteor>
- 2020-2021 **PhD student mentoring (MIT)**  
Mentored and collaborated with three PhD students in various research projects at Aleksander Mądry's lab at MIT; guided them through the research process, paper writing, and publication cycle; has resulted in two submissions that are under review at machine learning conferences.
- 2021 **Graduate Application Assistance Program**  
Assisted applicants from under-represented groups with their graduate student applications to MIT's EECS PhD program.  
<https://www.thrive-eeecs.mit.edu/gaap>
- 2020-2021 **Masters student mentoring (CMU)**  
Directly supervised a Masters student in the Robotics Institute at CMU; lead student through near completion of a research project, resulting in a conference paper currently under submission; the student is now a PhD student at the Computer Science Department at Stanford University.
- 2019-2020 **PhD student mentoring (CMU)**  
Mentored and collaborated with a first-year PhD student in the Computer Science Department at CMU; guided the student through the entire publication process, resulting in two conference papers published ICLR and ICML.
- 2019-2020 **CMU AI Mentoring Program**  
Mentored undergraduate women and minorities in one-on-one meetings to provide career advice and discuss research/graduate school; the mentee is now a PhD student at UC Berkeley.
- 2019-2020 **Teknowledge at Obama Academy**  
Taught middle schoolers how to code as part of the Teknowledge outreach program at the Obama Academy; courses were intended to provide early exposure to computer science for under-represented students in low-income neighborhoods of Pittsburgh
- 2019 **Mental Health First Aid Certification**  
Underwent training to recognize mental health issues and provide first aid assistance to those in need  
<https://www.mentalhealthfirstaid.org/>

2018-2020      **Undergraduate student mentoring (CMU)**  
Directly supervised a visiting undergraduate from IIT Delhi on a research project at CMU; continued guiding and supporting the student throughout the publication cycle culminating in the completion of the project and an ICML 2020 publication; the student is now a PhD student in the Machine Learning Department at Carnegie Mellon University.

## Professional service

---

2022      **Workshop on Adversarial Machine Learning and Beyond**  
Organizer for an AAAI 2022 workshop broadly themed around adversarial machine learning  
Website: <https://advml-workshop.github.io/aaai2022/>

2021      **A Blessing in Disguise: The Prospects and Perils of Adversarial Machine Learning**  
Organizer for an ICML 2021 workshop themed around the dangers and benefits of adversarial machine learning.  
Website: <https://advml-workshop.github.io/icml2021/>

2021      **Robust and reliable ML in the real world**  
Main organizer for an ICLR 2021 workshop on real world robustness.  
Website: <https://sites.google.com/connect.hku.hk/robustml-2021/home>

## Awards

---

2020      **SCS Dissertation Award — Honorable Mention**, *Provable, structured, and efficient methods for robustness of deep networks to adversarial examples*, Carnegie Mellon University

2020      **Siebel Scholar Fellowship**, Carnegie Mellon University

2017      **Best Defense Paper at NeurIPS 2017 ML & Security Workshop**, *Provable defenses against adversarial examples via the convex outer adversarial polytope*

2013      **Summer Undergraduate Research Fellowship**, Carnegie Mellon University

## All publications

---

2021      **Under review**      **Missingness bias in model debugging**  
Saachi Jain, Hadi Salman, Eric Wong, and Aleksander Madry

arXiv      **Certified patch robustness via smoothed vision transformers**  
Hadi Salman, Saachi Jain, Eric Wong, and Aleksander Madry  
*arXiv:2110.07719 [cs.CV]*

arXiv      **DeepSplit: Scalable verification of deep neural networks via operator splitting**  
Shaoru Chen, Eric Wong, J Zico Kolter, and Mahyar Fazlyab  
*arXiv:2106.09117 [cs.LG]*

ICML      **Leveraging sparse linear layers for debuggable deep networks**  
Eric Wong, Shibani Santurkar, and Aleksander Madry  
*International Conference on Machine Learning*

- ICLR **Learning perturbation sets for robust machine learning**  
Eric Wong and J Zico Kolter  
*International Conference on Learning Representations*
- 2020 ICML **Adversarial robustness against the union of multiple perturbation models**  
Pratyush Maini, Eric Wong, and Zico Kolter  
*International Conference on Machine Learning*
- DiffCVGP **Semantic Adversarial Robustness with Differentiable Ray-Tracing**  
Rahul Venkatesh, Eric Wong, and J Zico Kolter  
*Workshop on Differentiable Vision, Graphics, and Physics in Machine Learning at NeurIPS 2020 (Workshop paper)*
- ICML **Overfitting in adversarially robust deep learning**  
Leslie Rice, Eric Wong, and Zico Kolter  
*International Conference on Machine Learning*
- ICLR **Fast is better than free: Revisiting adversarial training**  
Eric Wong, Leslie Rice, and J. Zico Kolter  
*International Conference on Learning Representations*
- IEEE IV **Neural network virtual sensors for fuel injection quantities with provable performance specifications**  
Eric Wong, Tim Schneider, Joerg Schmitt, Frank R Schmidt, and J Zico Kolter  
*2020 IEEE Intelligent Vehicles Symposium (IV)*
- 2019 ICML **Wasserstein adversarial examples via projected sinkhorn iterations**  
Eric Wong, Frank Schmidt, and Zico Kolter  
*International Conference on Machine Learning*
- 2018 NeurIPS **Scaling provable adversarial defenses**  
Eric Wong, Frank Schmidt, Jan Hendrik Metzen, and J. Zico Kolter  
*Advances in Neural Information Processing Systems*
- ICML **Provable defenses against adversarial examples via the convex outer adversarial polytope**  
Eric Wong and Zico Kolter  
*International Conference on Machine Learning*
- 2017 ICML **A semismooth Newton method for fast, generic convex programming**  
Alnur Ali, Eric Wong, and J Zico Kolter  
*International Conference on Machine Learning*
- 2015 AAAI **An SVD and derivative kernel approach to learning from geometric data**  
Eric Wong and J Zico Kolter  
*Twenty-Ninth AAAI Conference on Artificial Intelligence*