Buyun Liang

byliang@seas.upenn.edu buyunliang.org Google Scholar GitHub

EDUCATION

University of Pennsylvania — **Philadelphia, PA, USA**
Ph.D. in Computer and Information Science | Advisor: Prof. René Vidal — *Sep 2023 - Present*
- GPA: 3.88/4.0

University of Minnesota, Twin Cities — **Minneapolis, MN, USA**
M.Sc in Computer Science | Advisor: Prof. Ju Sun — *Sep 2020 - May 2023*
- GPA: 4.0/4.0

University of Minnesota, Twin Cities — **Minneapolis, MN, USA**
M.Sc in Materials Science (Ph.D. Track) | Advisor: Prof. Ilja Siepmann — *Sep 2018 - Aug 2020*
- GPA: 3.66/4.0 | GPA of AI-related courses: 4.0/4.0

Nanjing University — **Nanjing, Jiangsu, China**
B.Sc in Physics (Elite Program) — *Sep 2014 - Jun 2018*
- GPA: 89.6/100

PUBLICATIONS

Trustworthy AI: Safe & Reliable LLMs

[1] **SECA: Semantically Equivalent and Coherent Attacks for Eliciting LLM Hallucinations**
Buyun Liang, *Liangzu Peng, Jinqi Luo, Darshan Thaker, Kwan Ho Ryan Chan, René Vidal*
NeurIPS 2025 [**paper**][**code**]

[2] **KDA: A Knowledge-Distilled Attacker for Generating Diverse Prompts to Jailbreak LLMs**
Buyun Liang, *Kwan Ho Ryan Chan, Darshan Thaker, Jinqi Luo, René Vidal*
In submission, 2025 [**paper**]

Optimization for AI

[3] **NCVX: A General-Purpose Optimization Solver for Constrained Machine and Deep Learning.**
Buyun Liang, *Tim Mitchell, Ju Sun.*
NeurIPS 2022 Workshop: Optimization for Machine Learning [**paper**][**website**][**code**][**talk**]

[4] **NCVX: A User-Friendly and Scalable Package for Nonconvex Optimization in Machine Learning.**
Buyun Liang, *Tim Mitchell, Ju Sun.*
arXiv preprint, 2021 [**paper**]

Trustworthy AI: Adversarial Robustness in Vision

[6] **Optimization and Optimizers for Adversarial Robustness**
Hengyue Liang, ***Buyun Liang***, *Le Peng, Ying Cui, Tim Mitchell, Ju Sun*
arXiv preprint, 2023 [**paper**]

[7] **Optimization for Robustness Evaluation beyond ℓ_p Metrics**
Hengyue Liang, ***Buyun Liang***, *Ying Cui, Tim Mitchell, Ju Sun*
ICASSP 2023 & NeurIPS 2022 Workshop: Optimization for Machine Learning [**paper**][**poster**]

[8] **Implications of Solution Patterns on Adversarial Robustness**
Hengyue Liang, ***Buyun Liang***, *Ju Sun, Ying Cui, Tim Mitchell*
CVPR 2023 Workshop: Art of Robustness [**paper**]

AI for Science & Scientific Computing

[9] **Neural Topology Optimization Based on Differential Programming with Principled Constrained Optimization**

Ryan de Vera, ***Buyun Liang***, *Binyao Guo, Qizhi He, Ju Sun*
ASME 2024 International Mechanical Engineering Congress and Exposition [**paper**]

[10] **Large-scale molecular dynamics simulations of bubble collapse in water: Effects of system size, water model, and nitrogen**
Jingyi L Chen, Jesse L Prelesnik, ***Buyun Liang****, Yangzesheng Sun, Mrugank Bhatt, Christopher Knight, Krishnan Mahesh, J Ilja Siepmann*
The Journal of Chemical Physics (JCP Editors' Choice 2023) [**paper**]

[11] **Effect of Non-Condensable Gas on the Thermophysical Properties of Bubbly Water and on Bubble Collapse Dynamics Probed by Molecular Simulations.**
J. Ilja Siepmann, Jingyi L Chen, ***Buyun Liang****, Krishnan Mahesh*
The 33rd Symposium on Naval Hydrodynamics, 2020 [**paper**]

TALKS

SECA: Semantically Equivalent and Coherent Attacks for Eliciting LLM Hallucination
Neural Information Processing Systems (NeurIPS) *2025*

Deep Learning with Nontrivial Constraints
SIAM International Conference on Data Mining *2023*

When Deep Learning Meets Nontrivial Constraints
Midwest Machine Learning Symposium *2023*

Toward Trustworthy AI — Robustness and Beyond
Midwest Machine Learning Symposium *2023*

When Deep Learning Meets Nontrivial Constraints
3M poster session *2023*

NCVX: A General-Purpose Optimization Solver for Constrained Machine and Deep Learning
NeurIPS Workshop: Optimization for Machine Learning *2022*

PROFESSIONAL SERVICE

Organizer

Supported the organization and coordination of oral, tutorial, and keynote poster sessions @ 2025 Conference on Lifelong Learning Agents

Led the organization of a 2-Hour Tutorial ***"Deep Learning with Nontrivial Constraints"***, @ 2023 SIAM International Conference on Data Mining

Reviewer

Neural Information Processing Systems (NeurIPS)	*2025*
International Conference on Learning Representations (ICLR)	*2025-2026*
International Conference on Acoustics, Speech, and Signal Processing (ICASSP)	*2024-2026*
Uncertainty in Artificial Intelligence (UAI)	*2023-2025*
International Joint Conference on Neural Networks (IJCNN)	*2025*
IEEE International Workshop on Machine Learning for Signal Processing (MLSP)	*2025-2026*
Artificial Intelligence and Statistics (AISTAT)	*2023*
NeurIPS 2023 Workshop: Optimization for Machine Learning (OPT)	*2023*

(Update: October 19, 2025)