

# Apurva Badithela

+1 (626) 267 4650  
✉ [ab5832@princeton.edu](mailto:ab5832@princeton.edu)  
🔗 [Google Scholar](#)  
🐙 [Github](#)  
🌐 [LinkedIn](#)  
🏠 [Website](#)

## Summary

Researcher and engineer with 6+ years of experience in building reliable learning-enabled robotic systems in manipulation and self-driving. Expertise includes evaluation, safety testing, verification and validation, applied statistics, robotics, formal methods, control and systems theory, game theory, and optimization. Led multiple successful projects delivering *scalable end-to-end evaluation* for state-of-the-art vision-language-action models (VLAs), *modular, safety-aware evaluation* of perception, and *closed-loop test generation* to stress-test high-level task reasoning. Direct experience with real2sim for robotic manipulation, hypothesis testing, non-asymptotic statistics, action-conditioned video generation models, diffusion policies, temporal logic, discrete event systems, assume-guarantee reasoning, and modular and hierarchical system architectures. Driven to build a virtuous evaluation-development cycle for safe, robust, and trustworthy robots.

## Education

- 2018–2024 **PhD, Control and Dynamical Systems,** *California Institute of Technology*  
Thesis: *Test and Evaluation of Autonomous Systems: Reactive Test Synthesis and Task-Relevant Evaluation of Perception* [PDF]  
Thesis Committee: Richard M. Murray (advisor), Tichakorn Wongpiromsarn, Mani Chandy, Joel Burdick, and Aaron Ames
- 2014–2018 **Bachelors in Aerospace Engineering and Mechanics,** *University of Minnesota, Twin-Cities*  
*Summa cum laude*, GPA: 3.93/4.00  
Thesis: *Exploiting Structure in Semidefinite Programming Problems with Applications to Robust Control*  
Advisor: Peter J. Seiler

## Experience

- 2024– **Presidential Postdoctoral Research Fellow,** *Princeton University*  
Advisor: Anirudha Majumdar  
Projects: Reliable and scalable policy evaluation, policy comparison, guiding data collection for imitation learning, action-conditioned video generation models and physics-based simulation for robot manipulation
- 2018–2024 **PhD Candidate,** *Caltech*  
Advisor: Richard M. Murray  
Projects: Closed-loop test-case generation for high-level reasoning in learning-enabled robots, and task-relevant, modular evaluation of perception with respect to system-level safety requirements
- 2021 **Autonomy Research Intern in Behavior Planning and Prediction,** *Motional*  
Host: Eric Wolff, Tung Phan-Minh  
Project: Counterexample Guided Repair of an Inverse Reinforcement Learning Planner
- 2017 **ICES Moncrief Summer Research Fellow,** *University of Texas, Austin*  
Host: Ufuk Topcu, Ivan Papusha  
Project: Sparse Matrix Methods for Real-time Model Predictive Control
- 2016 **Summer Undergraduate Research Fellowship,** *Purdue University*  
Host: Neera Jain, Austin L. Nash  
Project: Dynamic Modeling and Validation of micro-CHP systems

---

## Selected Projects

- 2024– **Unreliable Simulators for Real-world Reliability,** *Princeton University*
- Developed a rigorous algorithmic framework for robot policy evaluation that combines large-scale simulation with small-scale real-world testing
  - Built a real-to-sim pipeline for robotic manipulation, supporting Franka and WidowX setups and leveraging ManiSkill, RoboCASA, Blender, and image-to-3D tools.
  - Reduced the number of hardware evaluations required by 20–25% while maintaining finite-sample statistical guarantees on real-world performance
  - Evaluated end-to-end visuomotor policies such as diffusion policy and foundation models such as  $\pi_0$
  - Managed a team of four PhD students and undergraduates
  - Led collaboration efforts with researchers at Waymo and UCLA
  - Received the Best Paper Award at CORL 2025 workshop on [EvalDeploy](#)
  - Investigating the application of action-conditioned video generation models for policy evaluation and safety-critical scenario generation
- 2025– **Statistically Rigorous Policy Evaluation and Comparison,** *Princeton University*
- Developed a statistically rigorous, sequential hypothesis testing framework to evaluate and compare policies with anytime stopping
  - Diagnosed and resolved hardware and software issues on the WidowX hardware platform, enabling sustained large-scale evaluation with over 500 hardware trials
  - Evaluated end-to-end visuomotor policies such as Octo and OpenVLA in both in-distribution and out-of-distribution environments, and analyzed evaluations from crowd-sourced platforms such as RoboArena
  - Demonstrated and proved Type-1 error control properties of proposed frameworks
  - Co-led collaborations with researchers at Toyota Research Institute (TRI)
  - Building robotics-focused algorithmic frameworks for efficient comparison of multiple policies in as few trials as possible with general task metrics
- 2021–2024 **Modular Evaluation of Object Detection in relation to System-level Safety,** *Caltech*
- Developed a formal framework to quantify the impact of perception errors and performance trade-offs (e.g., precision-recall) on system-level safety using confusion matrices
  - Proposed object-detection evaluation metrics that provide tighter bounds on system-level performance than standard measures
  - Derived perception-level performance requirements from system-level specifications using assume-guarantee reasoning
  - Demonstrated results on state-of-the-art object detection models using the nuScenes dataset
  - Contributed to open source Python toolbox [Pacti](#) for compositional system design and analysis in collaboration with researchers at UC Berkeley and NASA Jet Propulsion Laboratory (JPL).
- 2019–2024 **Closed-loop Test Generation for High-level Autonomy,** *Caltech*
- Developed novel frameworks for compositional and closed-loop testing of high-level planning and decision-making, evaluating robot behavior under feedback from its environment
  - Automated synthesis of test environments with both static obstacles and dynamic agents for evaluating high-level autonomy on quadrupeds in robotic navigation settings
  - Derived game-theoretic and network-flow based approaches to build test environment in closed-loop with system behavior while remaining agnostic to the robot policy
  - Implemented optimization routines such as mixed integer linear programs (MILPs), bilinear programs, min-max games with coupled constraints to effectively tackle combinatorial problem complexity

---

## Honors and Awards

- RSS Pioneer**
- 2026 Robotics: Science and Systems *Robotics: Science and Systems*  
Awarded to 30 early-career researchers per year worldwide
- Presidential Postdoctoral Research Fellowship**
- 2024– Highest university-wide postdoctoral fellowship award across all disciplines *Princeton University*  
Awarded to 12 postdocs per year worldwide
- Best Paper Award**
- 2025 Workshop on Evaluation and Deployment Across the Robot Learning Lifecycle *Eval&Deploy@CoRL2025*  
Conference on Robot Learning (CoRL) 2025
- CMS and IST Gradient for Change**
- 2022 Department award for contributions toward Caltech graduate experience *Caltech*

|         |                                                                                                           |                               |
|---------|-----------------------------------------------------------------------------------------------------------|-------------------------------|
| 2022    | <b>CMS TA Fellow</b><br>EAS Division award to support CMS department TAs in promoting inclusive learning. | Caltech                       |
| 2022    | <b>RSS Inclusion Fellow</b><br>Conference Award                                                           | Robotics: Science and Systems |
| 2018    | <b>Guidance, Navigation, and Control Undergraduate Conference Experience Award</b><br>Conference Award    | AIAA                          |
| 2016–18 | <b>Robert and John McCollum Scholarship</b><br>Department Award                                           | University of Minnesota       |
| 2014–18 | <b>Dean's List</b><br>School of Engineering                                                               | University of Minnesota       |
| 2014–18 | <b>Gold Global Excellence Scholarship</b><br>University-wide Award                                        | University of Minnesota       |

## Publications

- [18] **Beyond Binary Success: Sample-Efficient and Statistically Rigorous Robot Policy Comparison**  
David Snyder, **Apurva Badithela**, Nikolai Matni, Anirudha Majumdar, Masha Itkina, Haruki Nishimura, George J. Pappas  
*Under Review at RSS 2026*
- [17] **PlayWorld: Learning Robot World Models from Autonomous Play**  
Tenny Yin, Zhiting Mei, Zhonghe Zheng, Miyu Yamane, David Wang, Jade Sceats, Samuel M Bateman, Lihan Zha, **Apurva Badithela**, Ola Shorinwa, Anirudha Majumdar  
*Under Review at RSS 2026*
- [16] **Video Models in Robotics: Applications, Challenges, Future Directions [PDF]**  
Zhiting Mei\*, Tenny Yin\*, Ola Shorinwa\*, **Apurva Badithela**, Zhonghe Zheng, Joseph Bruno, Madison Bland, Lihan Zha, Asher Hancock, Jaime Fernández Fisac, Philip Dames, Anirudha Majumdar  
*Under Review at ACM Computing Surveys*
- [15] **Reliable and Scalable Robot Policy Evaluation with Imperfect Simulators [PDF][Project Page]**  
**Apurva Badithela**, David Snyder\*, Lihan Zha\*, Joseph Mikhail, Matthew O'Kelly†, Anushri Dixit†, Anirudha Majumdar  
**Best Paper Award** at *Workshop on Eval&Deploy: Evaluation and Deployment Across the Robot Learning Life-cycle, Conference on Robot Learning (CoRL 2025)*  
**Oral Paper** at *SAFE-ROL: 2nd Workshop on Safe and Robust Robot Learning for Operation in the Real World, Conference on Robot Learning (CoRL 2025)*  
*IEEE International Conference on Robotics and Automation, 2026.*
- [14] **Guiding Data Collection via Factored Scaling Curves [PDF] [Project Page]**  
Lihan Zha, **Apurva Badithela**, Michael Zhang, Justin Lidard, Jeremy Bao, Emily Zhou, David Snyder, Allen Z. Ren, Dhruv Shah, Anirudha Majumdar  
*Under Review.*
- [13] **Is Your Imitation Learning Policy Better than Mine? Policy Comparison with Near-Optimal Stopping [PDF] [Project Page]**  
David Snyder, Asher J. Hancock, **Apurva Badithela**, Emma Dixon, Patrick Miller, Rares Andrei Ambrus, Anirudha Majumdar, Masha Itkina, and Haruki Nishimura  
*Robotics: Science and Systems (RSS) 2025.*
- [12] **Flow-Based Synthesis of Reactive Tests for Discrete Decision-Making Systems with Temporal Logic Specifications [PDF][Tool]**  
Josefine B. Graebener\*, **Apurva S. Badithela\***, Denizalp Goktas, Wyatt Ubellacker, Eric V. Mazumdar, Aaron D. Ames, and Richard M. Murray  
*IEEE Open Journal of Control Systems (OJ-CSYS), 2025.*
- [11] **Task-Relevant Evaluation Metrics of Object Detection for Quantitative System-Level Analysis of Safety-Critical Autonomous Systems [PDF] [Code]**  
**Apurva Badithela**, Ranai Srivastav, Tichakorn Wongpiromsarn, and Richard M. Murray.  
*ACM Transactions on Cyber-Physical Systems (T-CPS): Special Issue on Embodied AI in Cyber-Physical Systems: Algorithms, Computing Systems, Applications, and Trustworthiness. To Appear.*

\* Denotes equal contribution, † denotes equal advising.

- [10] **Pacti: Scaling Assume-Guarantee Reasoning for System Analysis and Design [PDF][Tool]**  
Inigo Incer, **Apurva Badithela**, Josefine Graebener, Piergiuseppe Mallozzi, Ayush Pandey, Sheng-Jung Yu, Albert Beneveniste, Benoit Caillud, Richard M. Murray, Alberto Sangiovanni-Vincentelli, and Sanjit Seshia.  
*ACM Transactions on Cyber-Physical Systems (T-CPS)*, 2025, pp 1-35.
- [9] **Evaluation Metrics of Object Detection for Quantitative System-Level Analysis of Safety-Critical Autonomous Systems [PDF]**  
**Apurva Badithela**, Tichakorn Wongpiromsarn, and Richard M. Murray.  
*IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023.  
*CPS-IoT Week Workshop on Perception for Safety-Critical Cyber-Physical Systems*, 2023.
- [8] **Reasoning over Test Specifications using Assume-Guarantee Contracts [PDF]**  
**Apurva Badithela\***, Josefine Graebener\*, Inigo Incer\*, and Richard M. Murray.  
*Proceedings of the 15<sup>th</sup> NASA Formal Methods (NFM)*, 2023, pp 278-294.
- [7] **Synthesizing Reactive Test Environments for Autonomous Systems: Testing Reach-Avoid Specifications with Multi-Commodity Flows [PDF]**  
**Apurva Badithela\***, Josefine Graebener\*, Wyatt Ubellacker, Eric V. Mazumdar, Aaron D. Ames, and Richard M. Murray.  
*IEEE International Conference on Robotics and Automation (ICRA)*, 2023.  
*Workshop on Envisioning an Infrastructure for Multi-Robot and Collaborative Autonomy Testing and Evaluation, Robotics: Science and Systems (RSS)*, 2022.
- [6] **Towards Better Test Coverage: Merging Unit Tests for Autonomous Systems [PDF]**  
**Apurva Badithela\***, Josefine Graebener\*, **Apurva Badithela\***, and Richard M. Murray.  
*Proceedings of the 14<sup>th</sup> NASA Formal Methods (NFM)*, 2022, pp 133-155.
- [5] **Leveraging Classification Metrics for Quantitative System-level Analysis of Temporal Logic Specifications [PDF]**  
**Apurva Badithela**, Tichakorn Wongpiromsarn, and Richard M. Murray.  
*60<sup>th</sup> IEEE Conference on Decision and Control (CDC)*, 2021.
- [4] **Lipschitz continuity of signal temporal logic robustness measures: Synthesizing control barrier functions from one expert demonstration [PDF]**  
Prithvi Akella\*, **Apurva Badithela\***, Richard M. Murray, Aaron D. Ames.  
*arXiv* 2023.
- [3] **Synthesis of static test environments for observing sequence-like behaviors in autonomous systems [PDF]**  
**Apurva Badithela** and Richard M. Murray  
*arXiv* 2021.
- [2] **Analysis of the Heavy-ball Algorithm using Integral Quadratic Constraints [PDF]**  
**Apurva Badithela** and Peter Seiler.  
*2019 American Control Conference (ACC)*.
- [1] **Dynamic Modeling of a Sensible Thermal Energy Storage Tank with an Immersed Coil Heat Exchanger under Three Operation Modes [PDF]**  
Austin Nash, **Apurva Badithela**, and Neera Jain.  
*Journal of Applied Energy*, 2017.

## Invited Talks

|          |                                                       |                                                  |
|----------|-------------------------------------------------------|--------------------------------------------------|
| Dec 2025 | Robotics Graduate Student Organization (RGSO) Seminar | University of Delaware                           |
| Oct 2025 | Robotics Science Seminar                              | Amazon Robotics                                  |
| Oct 2025 | Princeton Robot Planning and Learning (PRPL) Lab      | Princeton University                             |
| Oct 2025 | Ames-Burdick Group Meeting                            | Caltech                                          |
| Oct 2025 | Dixit Lab                                             | UCLA                                             |
| Oct 2025 | Learning and Interactive Robot Autonomy Lab           | USC                                              |
| Oct 2025 | Waymo Reading Group Seminar                           | Waymo                                            |
| Dec 2023 | Toyota Motor North America R&D                        | Toyota Research Institute, North America (TRINA) |
| Nov 2023 | Autonomous Systems Lab (ASL) Group                    | Stanford University                              |
| Nov 2023 | ECE Department Seminar                                | University of Michigan, Ann Arbor                |
| Nov 2023 | Intelligent Robot Motion (IRoM) Lab                   | Princeton University                             |
| Oct 2023 | Ozay and Seiler Groups                                | University of Michigan, Ann Arbor                |

|          |                                                                   |                                          |
|----------|-------------------------------------------------------------------|------------------------------------------|
| Dec 2022 | Hasuo Lab                                                         | National Institute of Informatics, Tokyo |
| Oct 2022 | 40 <sup>th</sup> Southern California Controls Workshop            | Caltech                                  |
| Oct 2022 | AFOSR US-Japan Seminar on Autonomy, AI, Robotics, and Informatics | Tokyo                                    |
| Mar 2022 | VeHiCAL Group Meeting                                             | UC Berkeley                              |
| Dec 2020 | VeHiCAL Group Meeting                                             | UC Berkeley                              |

## Academic Service

### Mentorship

|         |                                                                           |                      |
|---------|---------------------------------------------------------------------------|----------------------|
| 2025–26 | Jade Sceaux — Reliable Evaluation with Video Models                       | Princeton University |
| 2025    | Joseph Mikhail — Reliable and Scalable Robot Policy Evaluation            | UT Austin            |
| 2024    | David Snyder — Hardware demonstrations of policy comparisons              | Princeton University |
| 2024    | Lihan Zha — Guided Data Collection with Factored Scaling Curves           | Princeton University |
| 2023–24 | Kimia Hassibi, Jacob Alderete - Test Generation                           | Caltech              |
| 2021–24 | Ranai Srivastav - Object Detection Metrics for Autonomous Vehicles        | Iowa State           |
| 2022    | Edward Zhang, Frida Moreno, Gerard Decker, Andy Dimnaku - Test Generation | Caltech              |
| 2020    | Berlin Del Aguila - Test Generation                                       | Caltech              |

### Teaching

|             |                                                         |         |
|-------------|---------------------------------------------------------|---------|
| Spring 2022 | Teaching Assistant, Optimal Control (CDS 112 / Ae 103a) | Caltech |
| Fall 2020   | Teaching Assistant, Linear Systems Theory (CDS 131)     | Caltech |
| Fall 2019   | Course Ombudsperson, Distributed Computing (CS 142)     | Caltech |

### Professional Service

|         |                                                                    |                      |
|---------|--------------------------------------------------------------------|----------------------|
| 2026    | Reviewer for Workshop Proposals, Robotics Science and Systems      | RSS                  |
| 2025–26 | Organizing Committee Member, Robotics Seminar                      | Princeton University |
| 2022–24 | Organizing Committee Member, H.B. Keller Colloquium Seminar        | Caltech              |
| 2021–22 | Member, Computing and Mathematical Sciences DEI Steering Committee | Caltech              |
| 2020–21 | Workshop Organizer, Building Effective Research Collaborations     | Caltech              |

### Review Activities

|         |                                                                                     |
|---------|-------------------------------------------------------------------------------------|
| 2026    | International Association of Safe and Ethical AI (IASEAI)                           |
| 2025–26 | Robotics: Science and Systems (RSS)                                                 |
| 2025–26 | IEEE Transactions on Robotics (T-RO)                                                |
| 2022–26 | IEEE International Conference on Robotics and Automation (ICRA)                     |
| 2023–25 | IEEE/RSJ Robotics and Automation Letters (R-AL)                                     |
| 2023–25 | IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)          |
| 2025    | ACM-IEEE International Conference on Cyber-Physical Systems (Poster and Demo Track) |
| 2024    | Nonlinear Analysis: Hybrid Systems (NAHS)                                           |
| 2023    | IEEE Transactions on Intelligent Transportation Systems (T-IST)                     |
| 2022    | Transactions on Automatic Control (TAC)                                             |
| 2021    | IEEE Conference on Decision and Control (CDC)                                       |

## Skills

|                             |                                                                                                                                                                                                                                                      |
|-----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Software</b>             | Python, PyTorch, ROS, Linux, Slurm, MATLAB                                                                                                                                                                                                           |
| <b>Robotics and Control</b> | Robot manipulation, policy evaluation, real-to-simulation, hardware–software integration, high-level autonomy and planning, automata theory, feedback control theory, system-theoretic analysis, assume-guarantee reasoning, safety, temporal logic. |
| <b>Statistics</b>           | Hypothesis testing, non-asymptotic statistics, confidence intervals, mean estimation, multiple testing                                                                                                                                               |
| <b>ML</b>                   | Diffusion policies, foundation models, action-conditioned video world models                                                                                                                                                                         |
| <b>Tools</b>                | ManiSkill, RoboCASA, OpenAI Gym, nuScenes                                                                                                                                                                                                            |