Pengcheng Wang

+1 (510)-282-5052 | wangpc@berkeley.edu | scholar.google.com/

Education

University of California, Berkeley — California, USA

Ph.D. Student — *Aug. 2024 - Present*

- Major: Control, Minors: Robotics, Learning and Optimization, GPA: 4.00
- Research Interests: Reinforcement Learning for Control, Dexterous Manipulation
- Research Advisor: Prof. Masayoshi Tomizuka (Member of the National Academy of Engineering)

Tsinghua University — Beijing, China

B. S. (major in Mechanics) and B. Eng. (major in Aerospace Engineering) — *Sep. 2020 - July 2024*

- GPA: 3.90, Rank 1^{st}
- Tang Lixin Scholarship (top 5% in 2021 for academic excellence)
- Academic Excellence Scholarship (top 3 in 2022 for academic excellence)

Working Experience

Offline Reinforcement Learning for Warehouse Manipulation — May. 2025 – Aug. 2025

Amazon Robotics — *Washington, USA*

- Develop a comlete pipeline of offline RL to improve upon rule-based expert for warehouse manipulation, achieving higher success rate with only binary labels of demonstrations

Research Experience

Domain Adaptive Diffusion Policy — Aug. 2025 – Present

Mechanical Systems Control Lab, UC Berkeley — *California, USA*

- Developed a novel diffusion policy with learned domain embedding, achieving SOTA performance on domain adaptation task compared with previous meta-RL methods
- By destroying the temporal dependencies in the training pairs, extract the static domain knowledge from context in an unsupervised approach with superior representative quality
- Instead of vanilla conditional diffusion, we introduce direct guidance on the diffusion process with the learned embedding, resulting in better performance and inference efficiency

Discrete Diffusion for Efficient VLA — Oct. 2025 – Present

Mechanical Systems Control Lab, UC Berkeley — *California, USA*

- Combine and extend the Real-time Chunking to Discrete Diffusion VLA for boosted efficiency and control quality
- Develop a discrete diffusion VLA upon a self-built preference-based dataset, greatly improving the language following ability of VLA models

Residual Q-Learning for Policy Customization — Aug. 2023 – Present

Mechanical Systems Control Lab, UC Berkeley — *California, USA*

- Bring the Residual Q-Learning (RQL) online, enabling test-time policy customization on super-human level racing agent GT Sophy 1.0 in the Gran Turismo Sport
- Apply the online-adaptation ability in LLM test-time-scaling, improve the model alignment performance without external reward models.
- Extend the RQL to policy gradient methods, introducing a unified framework for on-policy customization, and establishing clear connections between existing RLHF approaches
- Extend the RQL for multiple priors, enabling scaled-up RL that ingests parallelized environments

Residual Q-Learning for Policy Customization — Aug. 2023 – Sept. 2024

Mechanical Systems Control Lab, UC Berkeley — *California, USA*

- Bring the Residual Q-Learning (RQL) online, enabling test-time policy customization on super-human level racing agent GT Sophy 1.0 in the Gran Turismo Sport
- Apply the online-adaptation ability in LLM test-time-scaling, improve the model alignment performance without external reward models.
- Extend the RQL to policy gradient methods, introducing a unified framework for on-policy customization, and establishing clear connections between existing RLHF approaches

- Extend the RQL for multiple priors, enabling scaled-up RL that ingests parallelized environments

Transferable Policy Learning Aug. 2024 – Present

Mechanical Systems Control Lab, UC Berkeley *California, USA*

- Safely transfer the racing policy across different dynamics via a model-based approach. Require only 40 laps of data to carry out the transfer without failure on the GT Sophy 1.0

Model-based Reinforcement Learning Aug. 2024 – Jan. 2025

Mechanical Systems Control Lab, UC Berkeley *California, USA*

- Address the *policy mismatch* issue in TD-MPC2 with policy constraints, achieving state-of-the-art performance across various continuous-control benchmarks against existing methods

Publications

Pengcheng Wang*, Chenran Li*, et al. Residual-MPPI: Online Policy Customization for Continuous Control. **ICLR 2025.**

Pengcheng Wang*, Qinghang Liu*, et al. DADP: Domain Adaptive Diffusion Policy. Under Review.

Pengcheng Wang, et al. Residual Policy Gradient: A Reward View of KL-regularized Objective. **ICRA 2025 Safe-VLM Workshop Spotlight**

Haotian Lin, **Pengcheng Wang**, et al. TD-M (PC) ^{2}: Improving Temporal Difference MPC Through Policy Constraint. **L4DC 2026**

Fuxiang Zhang, **Pengcheng Wang**, et al. REAR: Scalable Test-time Preference Realignment through Reward Decomposition. Under Review

Guojian Zhan, **Pengcheng Wang**, et al. Mind Your Entropy: From Maximum Entropy to Trajectory Entropy-Constrained RL. Under Review

Guojian Zhan, Letian Tao, **Pengcheng Wang**, et al. Mean Flow Policy with Instantaneous Velocity Constraint for One-step Action Generation. **ICLR 2026 (Oral)**

Jianglan Wei*, Zhenyu Zhang*, **Pengcheng Wang***, et al. HD3C: Efficient Medical Data Classification for Embedded Devices. Under Review