Guojian Zhan

School of Vehicle and Mobility & College of AI
Tsinghua University
Phone: +86 18811385676
Email: zgj21@mails.tsinghua.edu.cn
Personal page: https://molumitu.github.io/

Research Overview

Born in 2000, I have come of age alongside the rapid evolution of artificial intelligence. My research is fundamentally rooted in Reinforcement Learning (RL), with a focus on integrating World Models and Generative AI to build self-evolving intelligent systems. I conceptualize World Models as predictive foundations for sample-efficient reasoning, while leveraging Generative AI to capture multi-modal, high-dimensional action distributions—a synergy crucial for the advancement of Vision-Language-Action (VLA) models. By grounding generative foundations and language intelligence within rigorous RL frameworks, I aim to develop agents capable of cross-modal understanding and self-correction, marking a definitive step toward Artificial General Intelligence (AGI).

Research Areas: Reinforcement Learning (RL).

Education

Ph.D. Tsinghua University, Beijing, CN **2021 – 2026**
SVM & College AI Advisor: Shengbo Eben Li

Visiting Ph.D. UC Berkeley, California, USA **2025 – 2026**
Berkeley AI Research (BAIR) Advisor: Masayoshi Tomizuka

B.S. Tsinghua University, Beijing, CN **2017 – 2021**
SVM Advisor: Shengbo Eben Li

Awards and Honors

ICLR Main Conference Oral Presentation (Top 1%) 2026
National Graduate Scholarship (Top 2%) 2025
Best Oral Award at IEEE CVCI (1/200+) 2023
Outstanding Graduate of Tsinghua University (Top 10%) 2021

Research Intern Experience

Didi Voyager Lab

Work on online RL post-training of multi-modal planning model.
Work on online RLVR for enhancing LLM reasoning capability with high training stability.

ByteDance Seed Robotics

Work on offline & online RL post-training of vision language models (VLM) for high-level planning.

Highlight Works

Mean Velocity Policy (MVP), ICLR 2026, Oral, Top 1%. This work marks the transition of Generative RL into the era of single-step inference, achieving high-fidelity generation and fast execution. We identify a theoretical limitation in current generative mean velocity field learning, the absence of boundary conditions, and consequently propose instantaneous state-wise velocity constraints with a rigorous proof of completeness. By integrating this approach with rejection sampling in an online RL setting, we developed MVP, which performs best across diverse robotic manipulation benchmarks.
It is promising to extend this efficient generative paradigm to VLA models, where the single-step inference capability of MVP can significantly reduce the latency of multimodal action tokens, enabling foundation models to achieve high-frequency, closed-loop control in complex physical environments.

Bootstrap Off-policy with World Model (BOOM), NeurIPS 2025. We propose BOOM, a novel world-model-based RL framework that establishes a bootstrap loop between the policy and the model-driven planner. Utilizing the policy to warm-start the planning process, the planner distills superior trajectories which, in turn, refine the policy via a likelihood-free alignment mechanism. BOOM achieves SOTA performance on the DeepMind Control Suite and HumanoidBench.

Stable RL for LLMs: A Sample-Trimming Approach, On-going. While using RL to improve LLM reasoning has great potential, the training is often unstable, showing sharp changes in entropy and huge jumps in gradients. We studied this problem at the sample level and found that removing just a tiny part of the samples can effectively stabilize the gradient norm and entropy change. Our results show that this simple method leads to steady and continuous performance growth during RL training.

Publications

Selected [RL] Papers

PB[1] **Guojian Zhan**, Letian Tao, Pengcheng Wang, Yixiao Wang, Yuxin Chen, Yiheng Li, Hongyang Li, Masayoshi Tomizuka, Shengbo Eben Li. "Mean Flow Policy with Instantaneous Velocity Constraint for One-step Action Generation". *International Conference on Learning Representations (ICLR)*, 2026. **[CCF-A conference, Oral, top 1%]**

PB[2] **Guojian Zhan**, Likun Wang, Feihong Zhang, Yang Guan, Shengbo Eben Li. "Harmonized Dual Policy Improvement for Model-based Reinforcement Learning". *International Conference in Machine Learning (ICML)*, 2026. **[CCF-A conference]**

PB[3] **Guojian Zhan**, Likun Wang, Xiangteng Zhang, Jiaxin Gao, Masayoshi Tomizuka, Shengbo Eben Li. "Bootstrap Off-policy with World Model". *Neural Information Processing Systems (NeurIPS)*, 2025. **[CCF-A conference]**

PB[4] Zhouyang Yu†, **Guojian Zhan†**, Yang Guan, Jingliang Duan, Letian Tao, Shengbo Eben Li. "Taming the Aleatoric Impulse in Off-Policy Reinforcement Learning". *International Conference in Machine Learning (ICML)*, 2026. **[CCF-A conference]**

PB[5] Likun Wang†, **Guojian Zhan†**, Shiqi Liu, Zeyu He, Yinuo Wang, Yatao Bian. "Wasserstein Generative Reinforcement Learning". *International Conference in Machine Learning (ICML)*, 2026. **[CCF-A conference]**

PB[6] Tianyi Zhang†, Likun Wang†, **Guojian Zhan†**, Feihong Zhang, Yang Guan, Yao Lyu, Shengbo Eben Li. "Langevin Rollout Optimization for Modelic Reinforcement Learning". *International Conference in Machine Learning (ICML)*, 2026. **[CCF-A conference]**

PB[7] **Guojian Zhan**, Jingliang Duanb, Yuxuan Jianga, Feihong Zhanga, Ziang Zhenga, Shengbo Eben Li. "Continuous-Time Policy Gradient". *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*. **[CCF-A journal, 2nd round revision]**

PB[8] Likun Wang, Xiangteng Zhang, Yinuo Wang, **Guojian Zhan**, Wenxuan Wang, Haoyu Gao, Jingliang Duan, Shengbo Eben Li. "Off-policy Reinforcement Learning with Model-based Exploration Augmentation". *Neural Information Processing Systems (NeurIPS)*, 2025. **[CCF-A conference]**

PB[9] **Guojian Zhan**, Xiangteng Zhang, Feihong Zhang, Letian Tao, Shengbo Eben Li. "Bicriteria Policy Optimization for High Accuracy Reinforcement Learning". *IEEE Transactions on Neural Networks and Learning Systems (TNNLS)*, 2025. **[CCF-B journal]**

PB[10] **Guojian Zhan**, Xin An, Yuxuan Jiang, Jingliang Duan, Huichan Zhao, Shengbo Eben Li. "Physics Informed Neural Pose Estimation for Real-Time Shape Reconstruction of Soft Continuum Robots". *IEEE Robotics and Automation Letters (RAL)*, 2025. **[CAAI-B journal]**

PB[11] **Guojian Zhan†**, Yuxuan Jiang†, Shengbo Eben Li, Yao Lyu, Xiangteng Zhang, Yuming Yin. "A Transformation-Aggregation Framework for State Representation of Autonomous Driving Systems". *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2025. **[CCF-B journal]**

PB[12] Yuxuan Jiang†, **Guojian Zhan†**, Zhiqian Lan, Chang Liu, Bo Cheng, Shengbo Eben Li. "A Reinforcement Learning Benchmark for Autonomous Driving in General Urban Scenarios". *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2023. **[CCF-B journal]**

PB[13] Ziang Zheng†, **Guojian Zhan†**, Bin Shuai, Shengtao Qin, Jiangtao Li, Tao Zhang, Shengbo Eben Li. "Transferable Latent-to-Latent Locomotion Policy for Efficient and Versatile Motion Control of Diverse Legged Robots". *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025. **[CAAI-B conference]**

PB[14] Yuheng Lei, Yao Lyu, **Guojian Zhan**, Tao Zhang, Jiangtao Li, Jianyu Chen, Shengbo Eben Li, Sifa Zheng. "Zeroth-Order Actor-Critic: An Evolutionary Framework for Sequential Decision Problems". *IEEE Transactions on Evolutionary Computation (TEVC)*, 2025. **[CCF-B/CAAI-A journal]**

PB[15] Yuxuan Jiang, Yujie Yang, Zhiqian Lan, **Guojian Zhan**, Shengbo Eben Li, Qi Sun, Jian Ma, Tianwen Yu, Changwu Zhang. "Rocket Landing Control with Random Annealing Jump Start Reinforcement Learning". *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024. **[CAAI-B conference]**

PB[16] Chen Chen, Zhiqian Lan, **Guojian Zhan**, Yao Lyu, Bingbing Nie, Shengbo Eben Li. "Quantifying the Individual Differences of Drivers' Risk Perception via Potential Damage Risk Model". *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2024. **[CCF-B journal]**

PB[17] Yangang Ren, Jianhua Jiang, **Guojian Zhan**, Shengbo Eben Li, Chen Chen, Keqiang Li, Jingliang Duan. "Self-learned Intelligence for Integrated Decision and Control of Automated Vehicles at Signalized Intersections". *IEEE Transactions on Intelligent Transportation Systems (TITS)*, 2022. **[CCF-B journal]**