UMBC

ICML
International Conference
On Machine Learning

Learning Anisotropic Value Geometry with Finsler Reinforcement Learning

ICML 2026

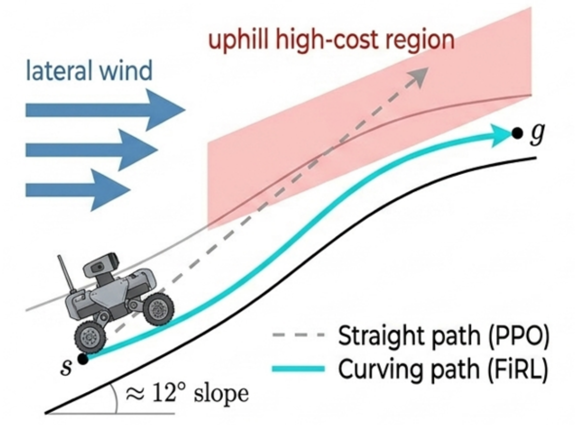
Jumman Hossain and Nirmalya Roy
University of Maryland, Baltimore County, USA

Motivation: Standard RL Struggles on Directional Terrain

Problem

Standard locomotion RL often treats costs as direction-agnostic and optimizes average return.

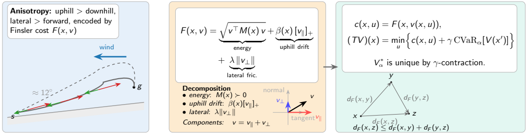
- Uphill motion is not the same as downhill motion.
- Lateral slip can be riskier than forward motion.
- Rare failures such as slipping, falling, or hard impacts matter more than average performance.
- A policy can look good on flat terrain but fail under slopes, wind, or contact-rich disturbances.



FiRL avoids the high-cost uphill region while PPO follows a more direct risky path.

FiRL: From Terrain Physics to Risk-Sensitive Control

- **Core Idea**: Integrate differential geometry with risk-sensitive RL for legged locomotion.
- **Finsler Geometry**: Introduces a state-and-velocity dependent metric $F(x, v)$ to model anisotropic effort.
- **Risk-Aversion**: Optimizes a **Conditional Value-at-Risk (** $CVaR_\alpha$ **)** objective to minimize the worst-case outcomes (the tail of the cost distribution).



FiRL Overview. Anisotropic terrain induces direction-dependent effort; the local Finsler cost encodes this structure, and the dynamic CVaR Bellman backup optimizes the tail-risk-aware policy.

Key Takeaway

The local cost geometry shapes the value landscape, while CVaR discourages rare high-cost outcomes.

Deconstructing the Finsler Cost Function

$$F(x, v) = w_e F_{\text{energy}}(x, v) + w_d F_{\text{drift}}(x, v) + w_f F_{\text{friction}}(x, v)$$

1. Kinetic Energy (F_{energy})

$$\sqrt{v^\top M(x) v}$$

Highlights the symmetric dynamic effort and state-dependent inertia matrix $M(x) \succ 0$.

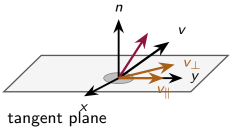
2. Uphill Drift Asymmetry (F_{drift})

$$\beta(x)[n(x)^\top v]_+$$

The $[\cdot]_+$ operator ensures only climbing against gravity $n(x)$ incurs the state-dependent penalty $\beta(x)$. This breaks spatial symmetry.

3. Lateral Friction (F_{friction}) $\lambda \| v_\perp \|$

Penalizes non-forward, lateral slipping $v_\perp$ to discourage high-curvature, low-traction trajectories.



Theory: CVaR–Finsler Bellman Operator

The Risk-Sensitive Backup: The optimal value function V_α^* is the fixed point of the operator T defined as:

$$(TV)(x) = \min_{u \in U(x)} \{ \underbrace{c(x,u)}_{\text{Finsler Cost}} + \gamma \underbrace{\rho_\alpha[V(X')]}_{\text{Tail Risk}} \}$$

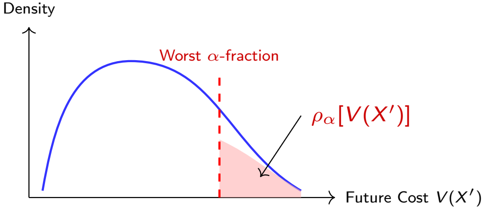
where $\rho_\alpha[\cdot]$ is the **Conditional Value-at-Risk ($CVaR_\alpha$)**.

Theorem 4.1: γ-Contraction Under mild boundedness conditions, T is a contraction mapping in the supremum norm:

$$\|TV_1 - TV_2\|_\infty \leq \gamma \|V_1 - V_2\|_\infty$$

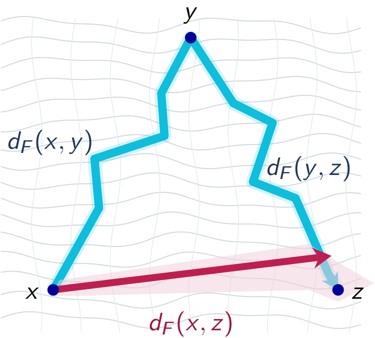
- **Unique Fixed Point**: Ensures the existence of a unique optimal value function V_α^*.

Tail-Risk Backup Visualization



Geometric Consistency:
- Immediate cost $c(x,u)$ is Finslerian.
- ρ_α biases the agent to avoid the "red zone" of catastrophic failure.

Induced Geometry and the Asymmetric Triangle Inequality



The Mathematical Insight

The local Finsler cost F induces an asymmetric global path cost:

$$\mathrm{d}_F(x, y) = \inf_{\tau : x \to y} \int_0^1 F(\tau(t), \dot{\tau}(t))\, dt$$

The Quasi-Metric Property

It allows $d_F(x, y) \neq d_F(y, x)$ because direction matters, while preserving the triangle inequality:

$$\mathrm{d}_F(x, z) \leq d_F(x, y) + d_F(y, z)$$

Key Takeaway

In Finsler geometry, the lowest-cost path on a slope need not be a straight line. It can be an energy-efficient tacking maneuver.

FiRL-AC Architecture

Environment / Physics Loop

- Agent takes action u_i in state x_i.
- Simulator computes the Finsler cost:

$$c_i = F(x_i, v(x_i, u_i)).$$

- Penalizes uphill and lateral motion.

Distributional Critic Loop

- Critic $Z_\phi(x)$ outputs K cost quantiles.
- Average the worst α-tail to estimate CVaR.
- Backup target:

$$y_i = c_i + \gamma \widehat{\mathrm{CVaR}}_\alpha.$$

Actor Update

- Policy is updated using a CVaR-based advantage:

$$\hat{A}_\alpha(x, u).$$

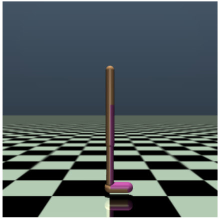
- The update discourages actions with high tail-risk outcomes.

Empirical Validation Across Environments

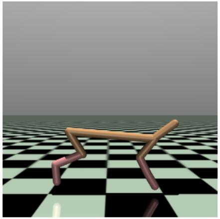
Controlled Stress Testing: MuJoCo

- **Tasks:** SlopedHopper-12°, Walker2d-5°, HalfCheetah-5°.
- **Purpose:** Isolate anisotropic effects such as incline, uphill effort, lateral drift, and wind disturbance.
- **Metrics:** Success rate, normalized energy, and $\mathrm{CVaR}_{0.1}$ tail cost.

Sloped Walker2d (5°)

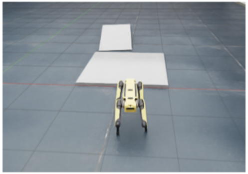


Sloped HalfCheetah (5°)



Realistic Contact Dynamics: Isaac Sim

- **Tasks:** Ramp Climb, Staircase, and Platform Beam traversal.
- **Purpose:** Evaluate contact-rich locomotion with a Spot-like quadruped under realistic 3D terrain interactions.
- **Metrics:** Success rate, energy per meter, tail cost, and fall rate.

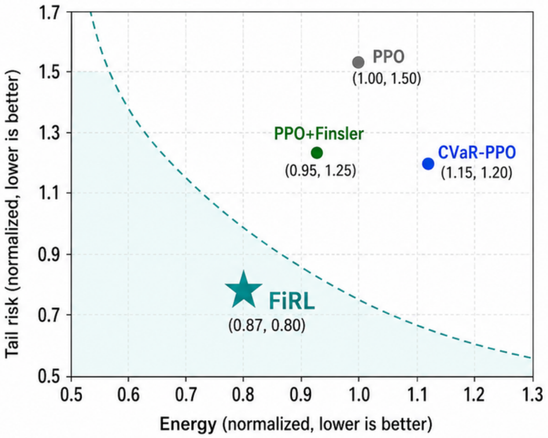


(a) Ramp climb

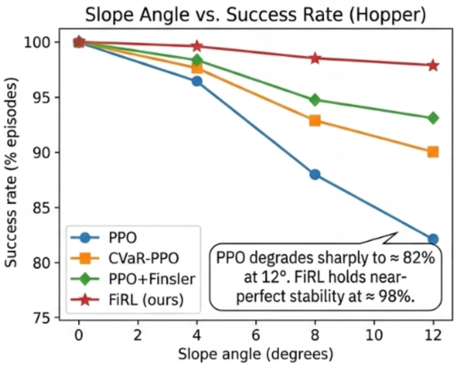
(b) Stair Case

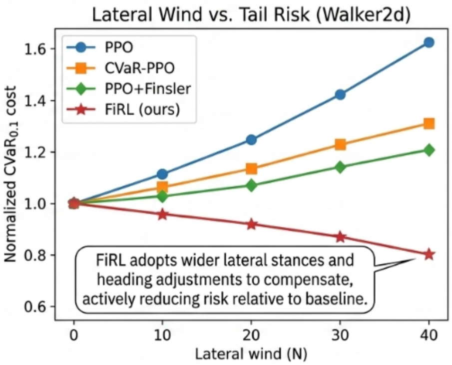
Results: Energy–Risk Pareto Dominance

- **Energy–risk tradeoff:** FiRL moves toward the bottom-left region, lowering both normalized energy and tail risk.
- **Performance:** On the 12° sloped Hopper task, FiRL reduces $\mathrm{CVaR}_{0.1}$ tail cost by about **35%** and total energy by about **15%** compared to PPO.
- **Behavioral shift:** CVaR-PPO reduces risk but can become overly conservative. FiRL instead uses the Finsler cost to prefer safer, lower-effort traversal directions.

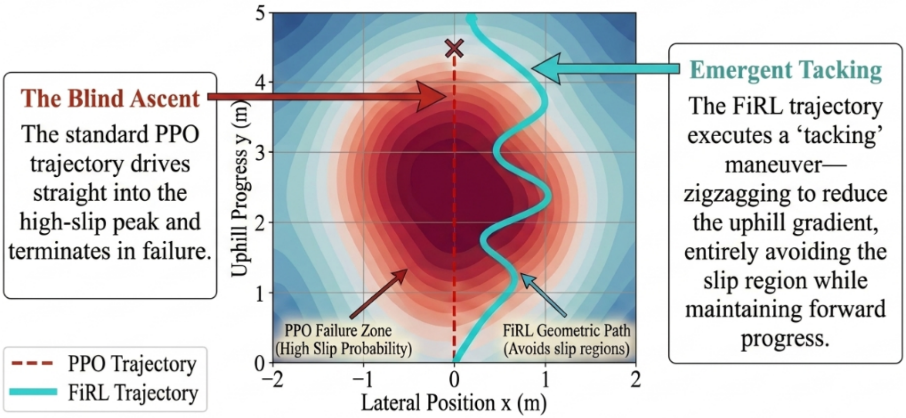


Results: Robustness Scaling



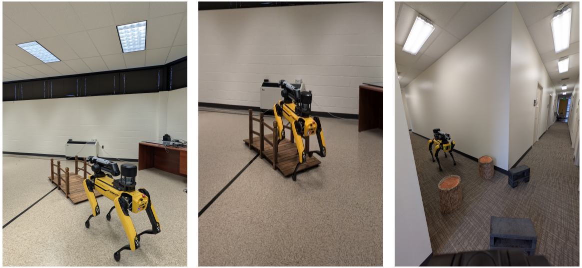


Behavioral Contrast: Risk vs Geometry



Physical-Robot Trials

- **Deployment:** We run the learned FiRL policy on a physical Boston Dynamics Spot robot.
- **Scenarios:** The robot is tested on narrow wooden supports and discrete foothold layouts.
- **Observation:** The policy produces cautious traversal and alignment behavior that is consistent with the simulation trends.



(a) Approach (b) Traversal (c) Footholds

Conclusion

- **Main idea:** FiRL combines a direction-dependent Finsler cost $F(x, v)$ with dynamic CVaR optimization for risk-sensitive legged locomotion.
- **Theory:** The CVaR–Finsler Bellman operator is a γ-contraction under bounded costs, and the induced path cost d_F gives an asymmetric quasi-metric geometry.
- **Results:** FiRL reduces tail risk and energy while maintaining higher success across MuJoCo, Isaac Sim, and preliminary physical-robot trials.
- **Future work:**
  - Adaptive tuning of anisotropy weights and risk level α.
  - End-to-end hardware deployment with onboard perception and repeated quantitative trials.

Thank You!

Questions and Discussion

Jumman Hossain

University of Maryland Baltimore County

jumman.hossain@umbc.edu

Project Page: pralgomathic.github.io/FiRL