Robust RL Navigation Via Sensitivity-Aware Observation Augmentation

James Zhu, Thierry Simeon, and Marco Cognetti

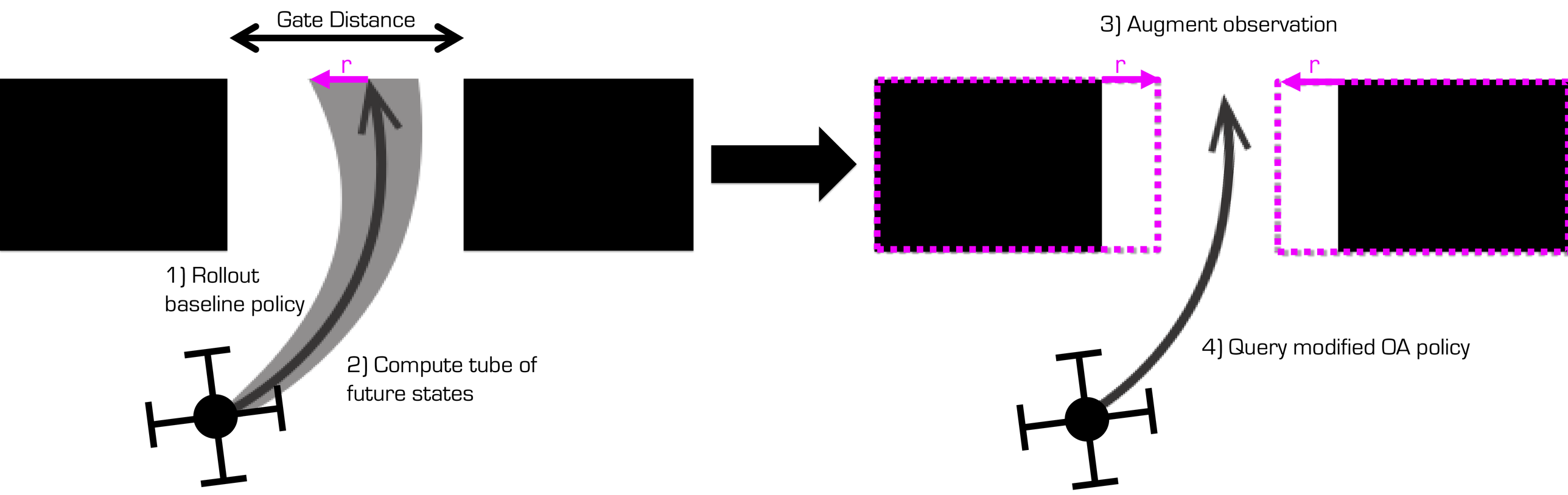


Figure 1: In order to be effective in real-world environments, learned policies must be able to handle estimation uncertainties. Obstacle augmentation (OA) leverages model-based sensitivity analysis to form a tube of future state deviations. With this information, the observation of the policy is modified to represent the "robustified" obstacle representation. This method allows us to improve policy robustness without increasing observation size or retraining.

Controlling Robots Under Uncertainties

- RL policies excel at robustly navigating uncertain environments, but don't explicitly quantify uncertainty.
- This can be a weakness when faced with sensing and estimation uncertainties at deployment.
- Incorporating uncertainty covariances into policy observations is challenging due to high dimensionality.
- Instead of adding hundreds of observation variables or retraining with additional domain randomization, we augment a baseline RL policy with sensitivity-aware chance-constraints, following [2].

Algorithm Outline (Figure 1)

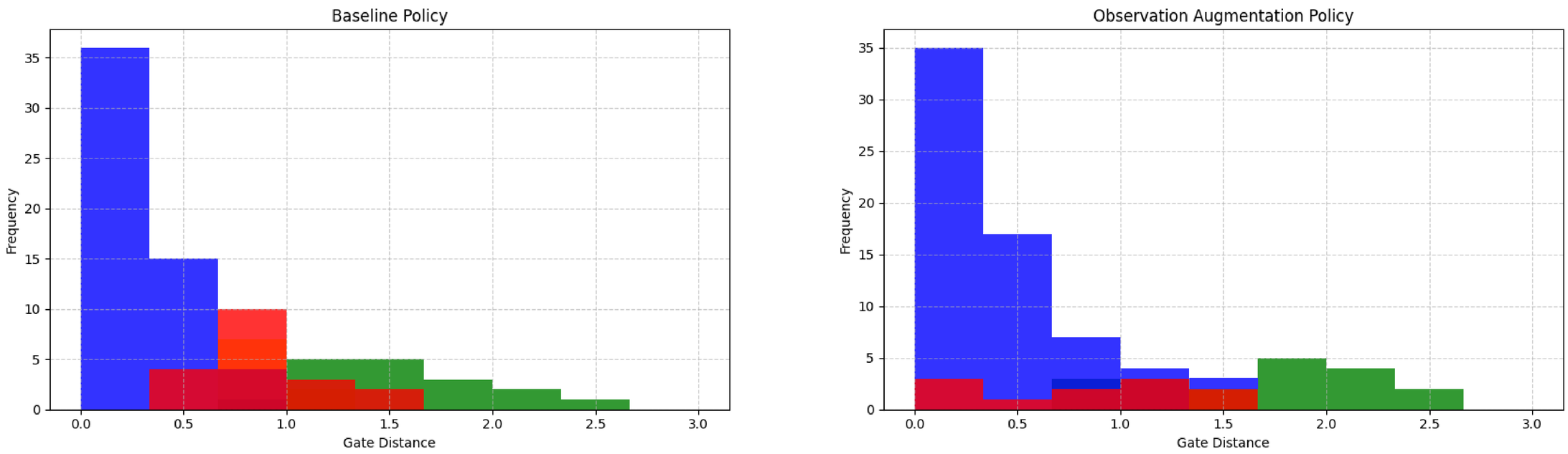
1. Rollout baseline policy
2. Compute tube of future states
3. Augment observation based on worst-case future state deviation
4. Query modified OA policy

Sensitivity-Aware Observation Augmentation

- Given a rollout of a baseline policy, sensitivity analysis [1] propagates uncertainty to form a "tube" of future states (Figure 1).
- This tube can be used to compute the worst-case future state deviation in the direction of the obstacles (r).
- We then modify the policy observation of the obstacle to account for the future state deviation.
- The ensuing policy will produce a safer motion that accounts for the effect of uncertainties.

Preliminary Results

- 100 trials of a gate passing task (Figure 1) were conducted with state, parameter, and obstacle estimation uncertainty drawn from an unscented Kalman filter.
- A baseline policy trained with PPO was compared to the same policy modified with observation augmentation.



Gate Successfully Passed:
Baseline: 21/100
Observation Augmentation: 14/100

Gate Not Passed:
Baseline: 60/100
Observation Augmentation: 75/100

Collision:
Baseline: 19/100
Observation Augmentation: 11/100

Figure 2: Histograms of 100 random trials for baseline policy and observation augmentation method. The baseline policy is unsafe when the gate distance is in a range where uncertainties in state and parameters can perturb the robot's trajectory so much that it results in a collision. The baseline policy experiences 19 obstacle failures (shown in red). Blue bars indicate when a robot does not pass through the gate but remains safe, and green bars are when the robot successfully passes through the gate. Using observation augmentation, the policy is more conservative, but more robust to uncertainties, experiencing only 11 collisions.

References: [1] A. Ansari and T. Murphey, "Minimum sensitivity control for planning with parametric and hybrid uncertainty," The Int. Journal of Robotics Research, vol. 35, no. 7, pp. 823–839, 2016.
[2] J. Zhu, T. Simeon, and M. Cognetti, "Robust sensitivity-aware chance-constrained MPC for efficient handling of multiple uncertainty sources," IEEE Robotics and Automation Letters, vol. 10, no. 10, pp. 10 330–10 337, 2025.

LAAS-CNRS
/ Laboratoire d'analyse et d'architecture des systèmes du CNRS

Laboratoire conventionné avec
Université de Toulouse