Realistic Adversarial Attacks for Robustness Evaluation of Trajectory Prediction Models via Future State Perturbation - Supplementary Information

Jeroen Hagenus, Julian F. Schumann, Frederik Baymler Mathiesen, Arkady Zgonnikov

I. Examples of generated attacks

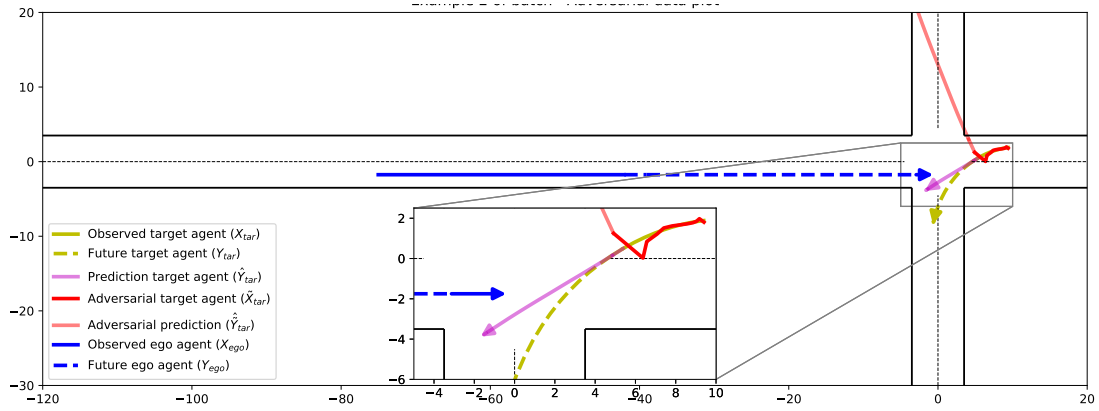

Fig. 1. Perturbing positions with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - ADE attack without dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed states.

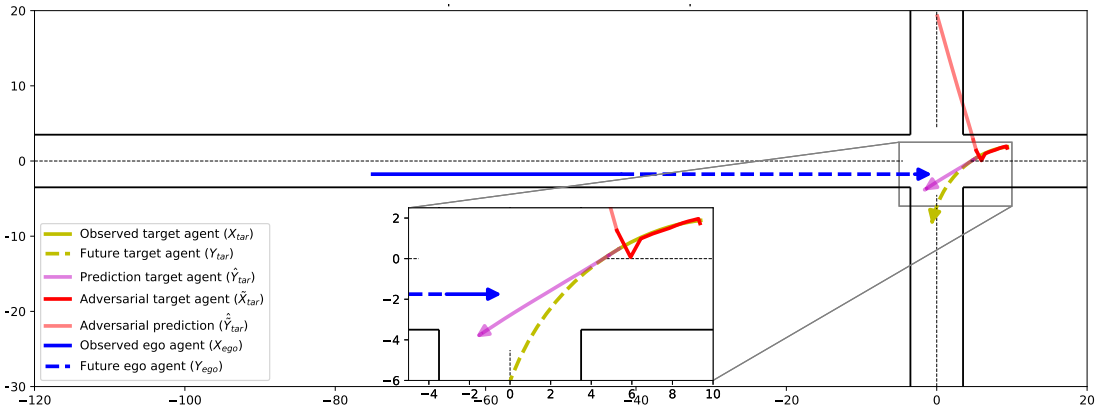

Fig. 2. Search attack with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - ADE attack using "Search" dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed states.

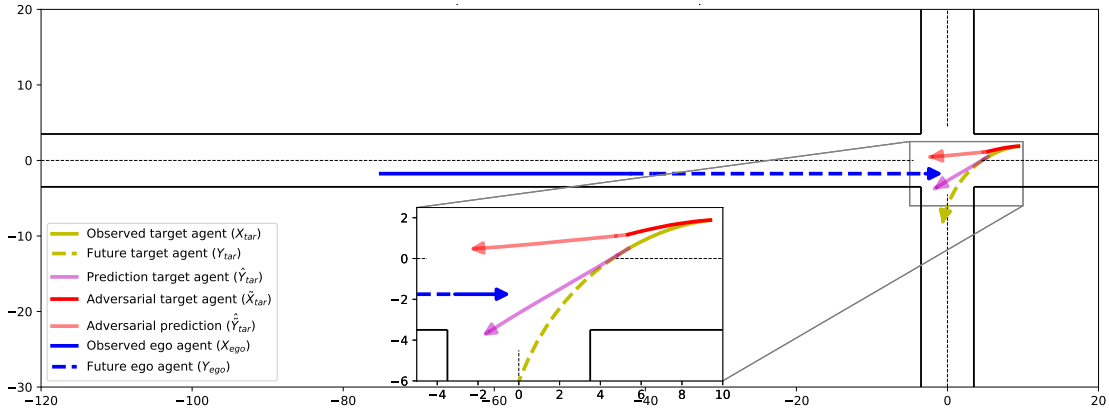

Fig. 3. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - ADE attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth.

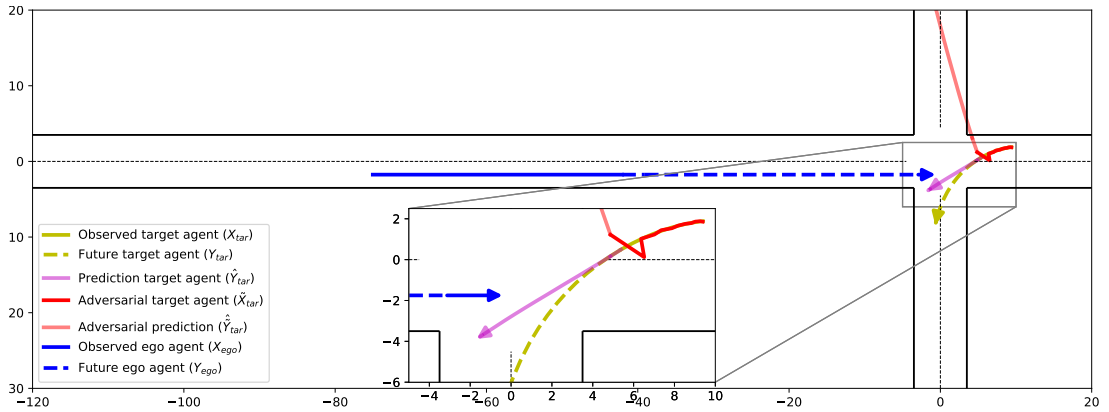

Fig. 4. Perturbing positions with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - FDE attack without dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed states.

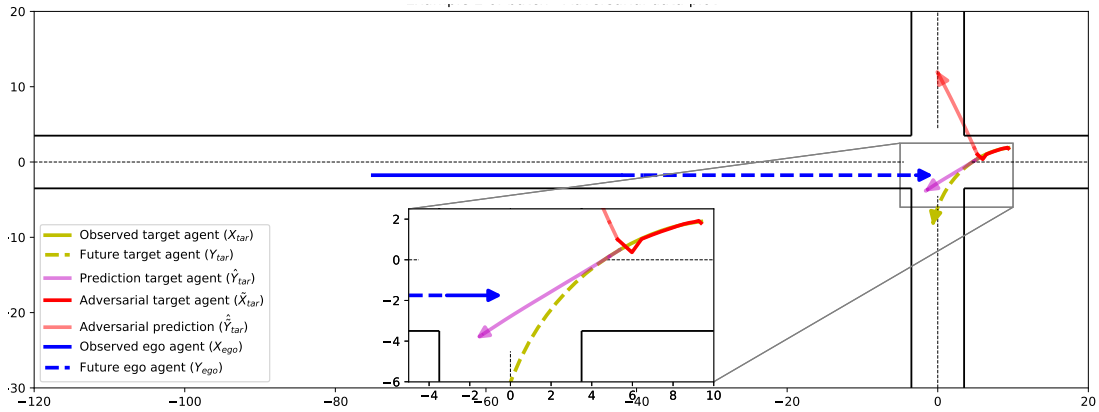

Fig. 5. Search attack with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - FDE attack with "Search" dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed states.

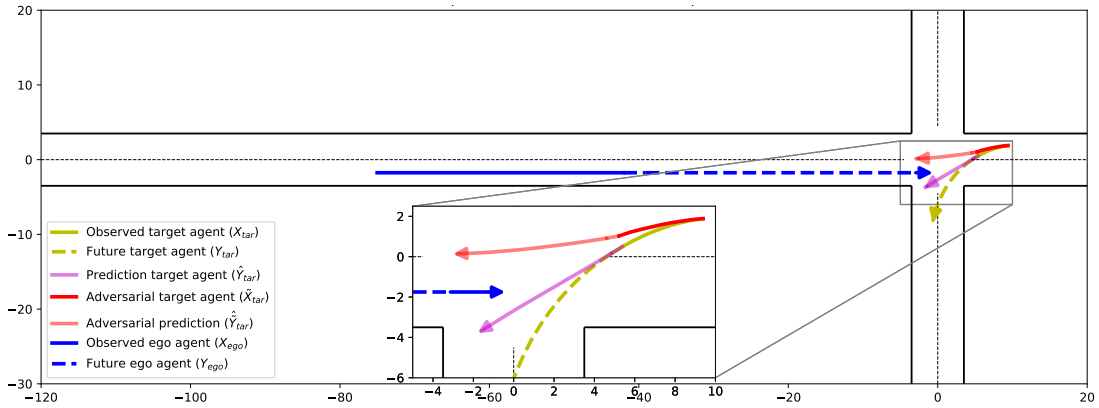

Fig. 6. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - FDE attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth.

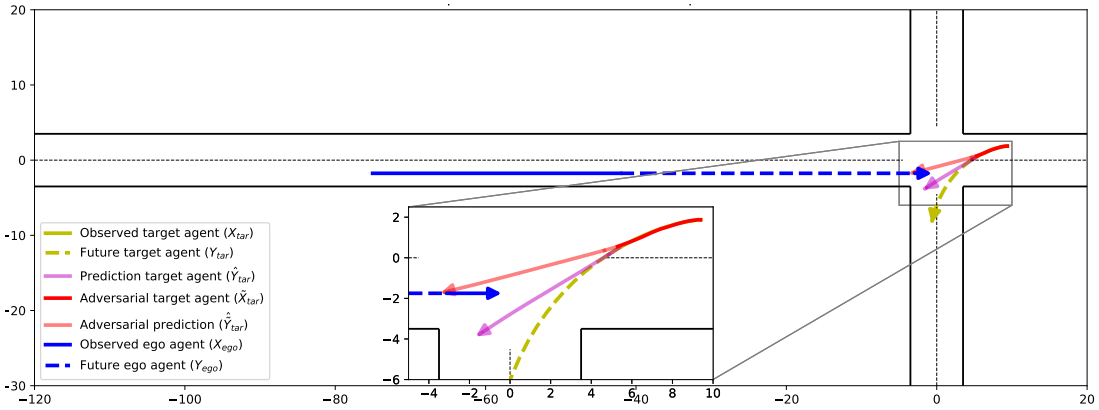

Fig. 7. Perturbing positions with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - False positive collision attack without dynamic constraints creates a realistic trajectory, with a small deviation in the final state of the observed states.

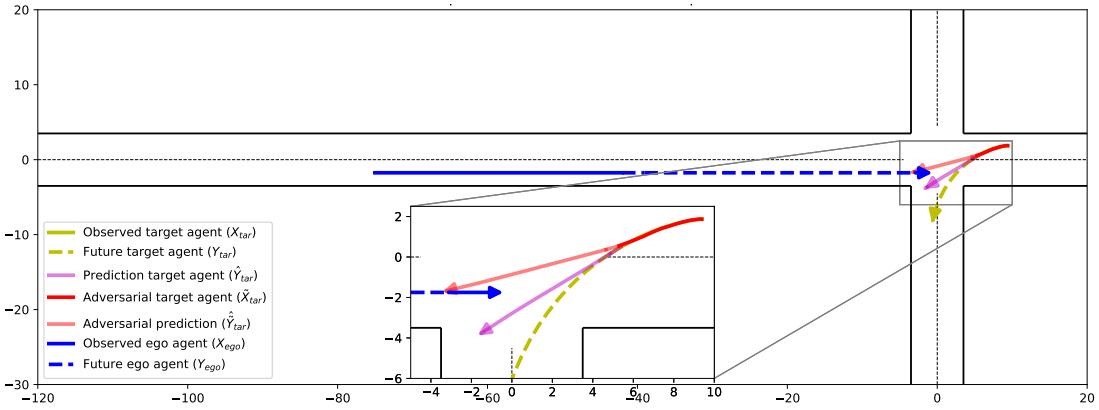

Fig. 8. Search attack with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - False positive collision attack with "Search" dynamic constraints creates a realistic trajectory, with a small deviation in the final state of the observed states.

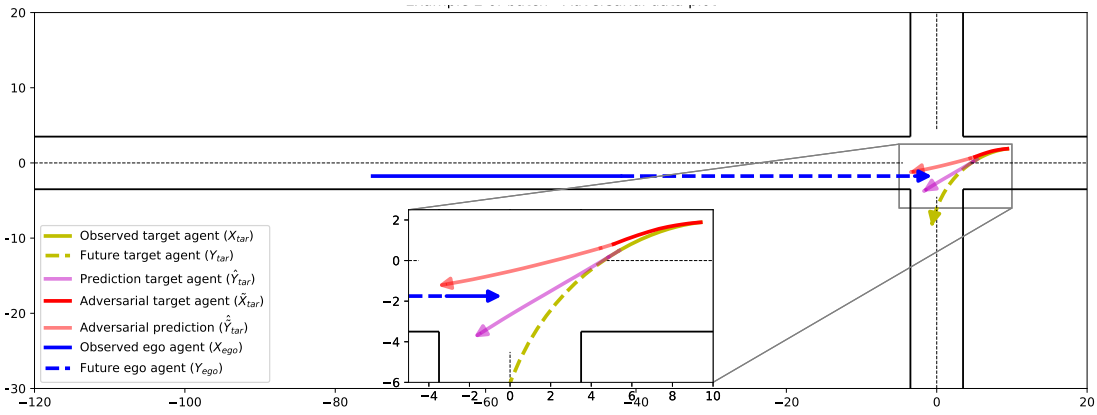

Fig. 9. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ - False positive collision attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth.

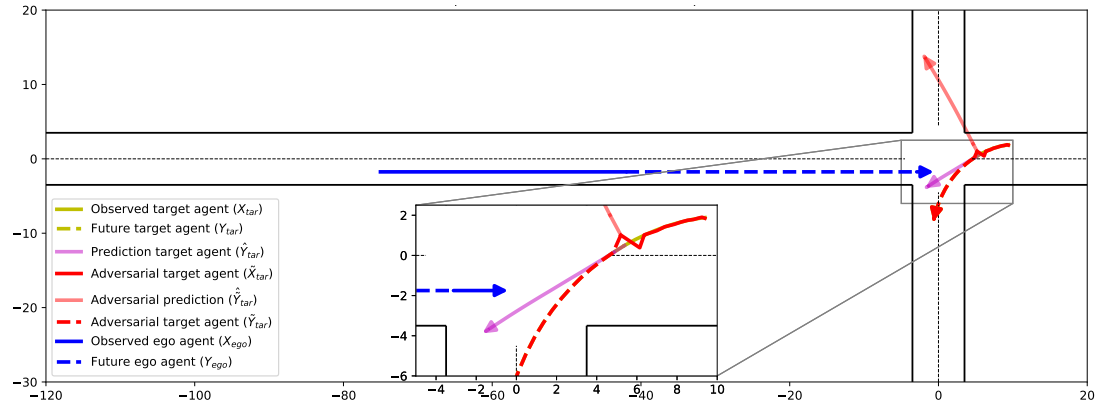

Fig. 10. Perturbing positions with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible ADE attack without dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed sequence. The final observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

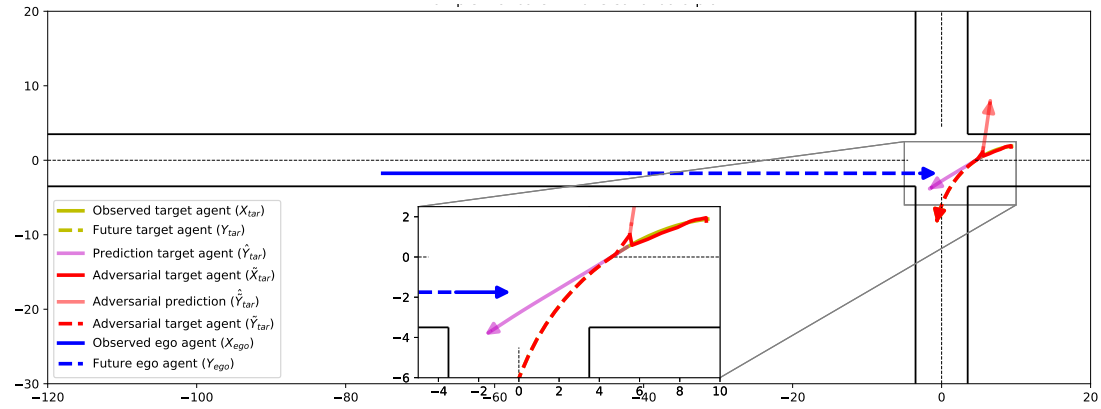

Fig. 11. Search attack with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible ADE attack with "Search" dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed sequence. The final adversarial observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

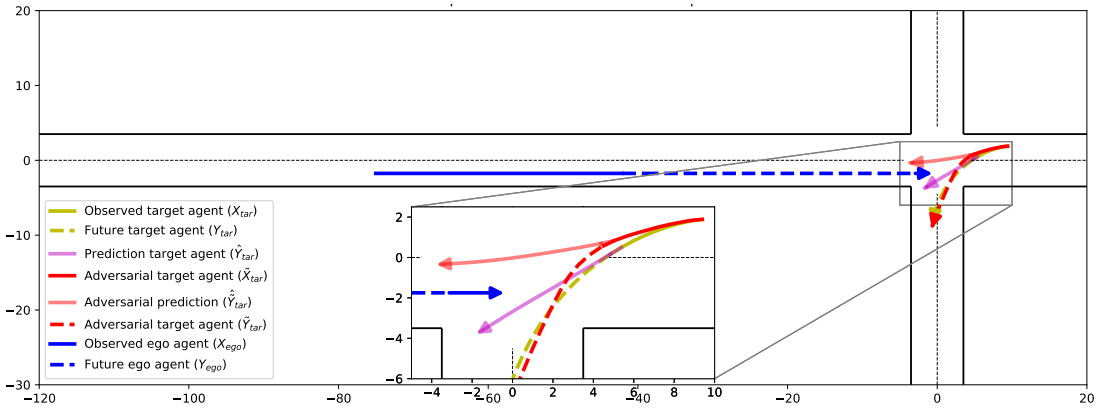

Fig. 12. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{ADE}}(Y_{\text{tar}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible ADE attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth. The adversarial observed states smoothly transition to the adversarial future states that closely resemble the future ground truth states.

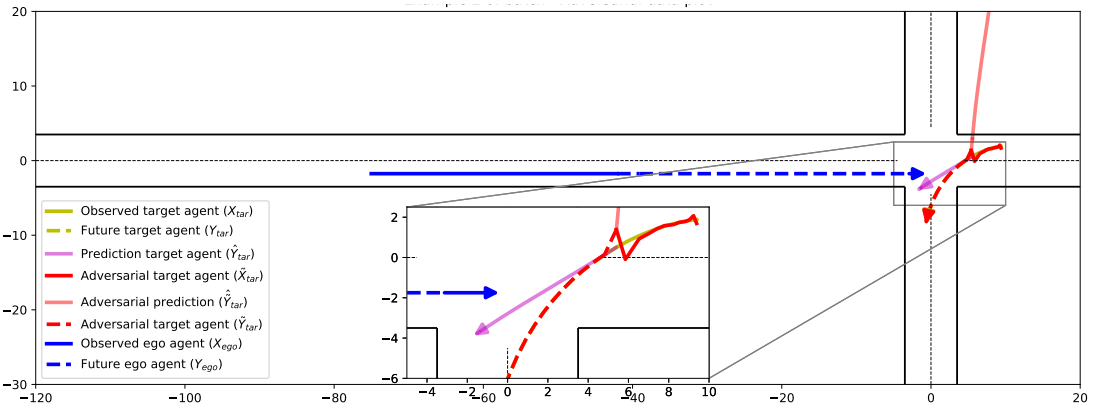

Fig. 13. Perturbing positions with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible FDE attack without dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed sequence. The final adversarial observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

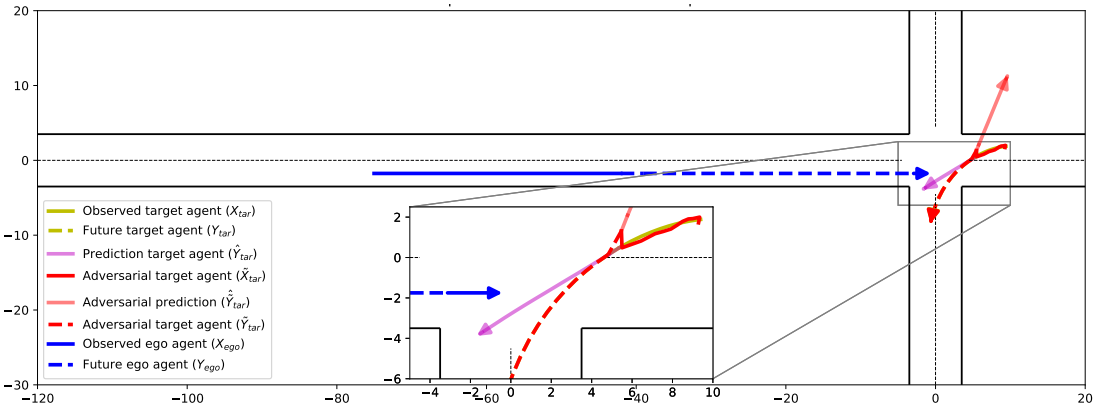

Fig. 14. Search attack with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible FDE attack with "Search" dynamic constraints creates an unrealistic trajectory, with large deviations in the final states of the observed sequence. The final adversarial observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

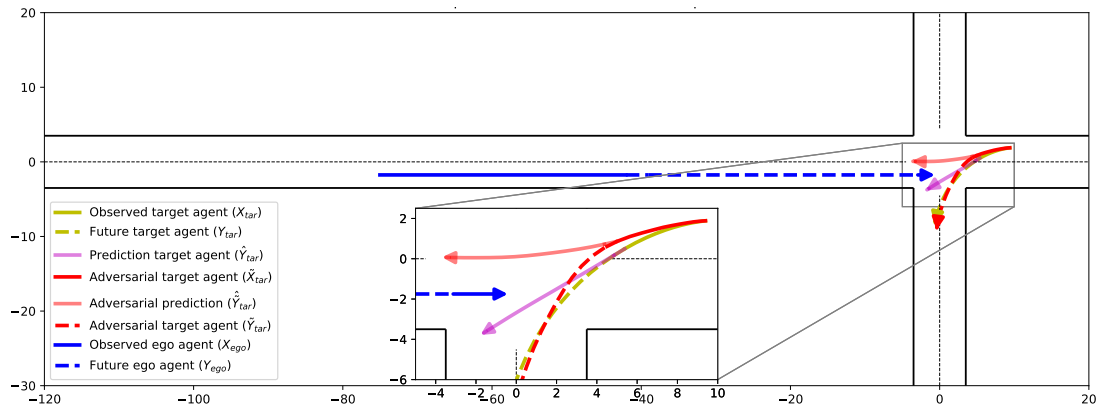

Fig. 15. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{FDE}}(Y_{\text{tar}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - Feasible FDE attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth. The adversarial observed states smoothly transition to the adversarial future states that closely resemble the future ground truth states.

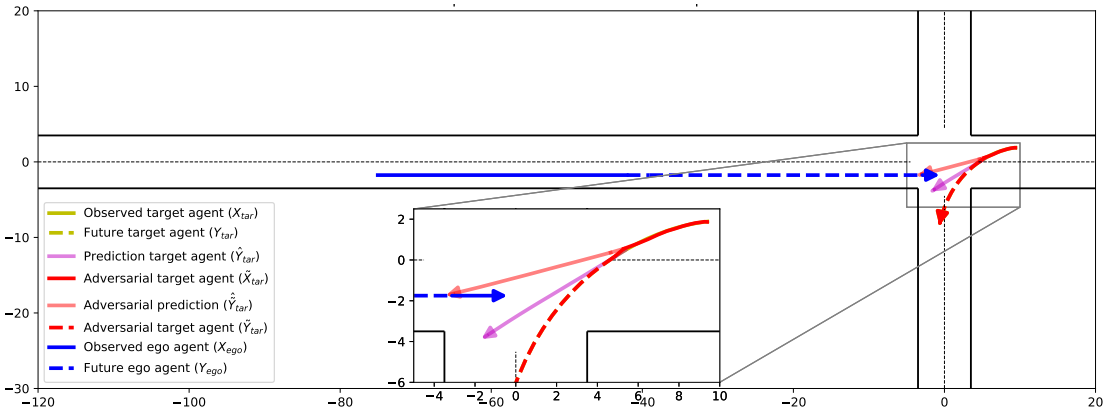

Fig. 16. Perturbing positions with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - False positive collision attack without dynamic constraints creates a realistic trajectory, with a small deviation in the final state of the observed states. The final adversarial observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

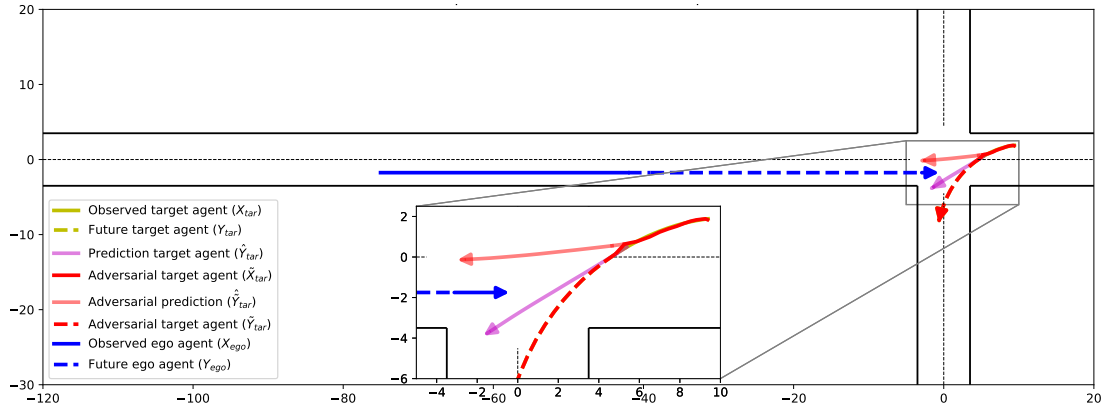

Fig. 17. Search attack with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \hat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - False positive collision with "Search" dynamic constraints creates a realistic trajectory, with a small deviation in the final state of the observed states. The final adversarial observed state is directly connected to the first state of the ground truth future states without modifications to those future states.

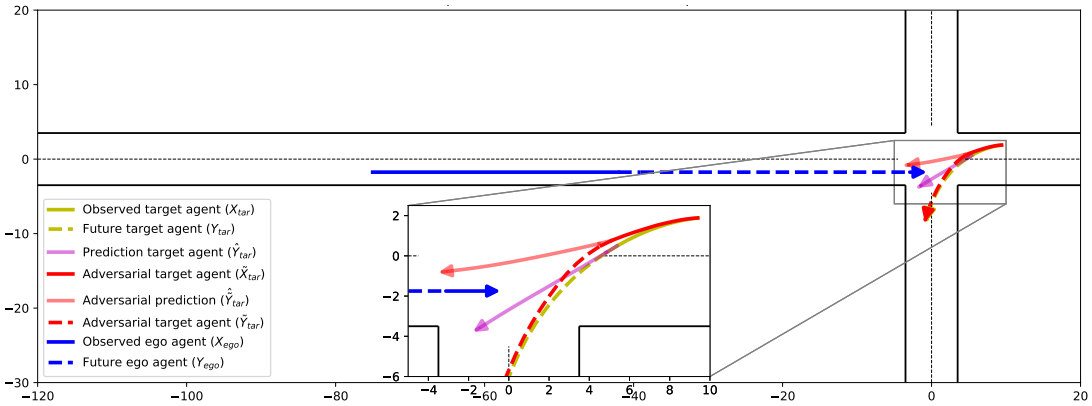

Fig. 18. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{Col}}(Y_{\text{ego}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}}) + l_{\text{Traj}}(\widetilde{Y}_{\text{tar}}, Y_{\text{tar}})$ - False positive collision attack with control action dynamic constraints creates a realistic trajectory, with a gradual increase in deviation from the ground truth. The adversarial observed states smoothly transition to the adversarial future states that closely resemble the future ground truth states.

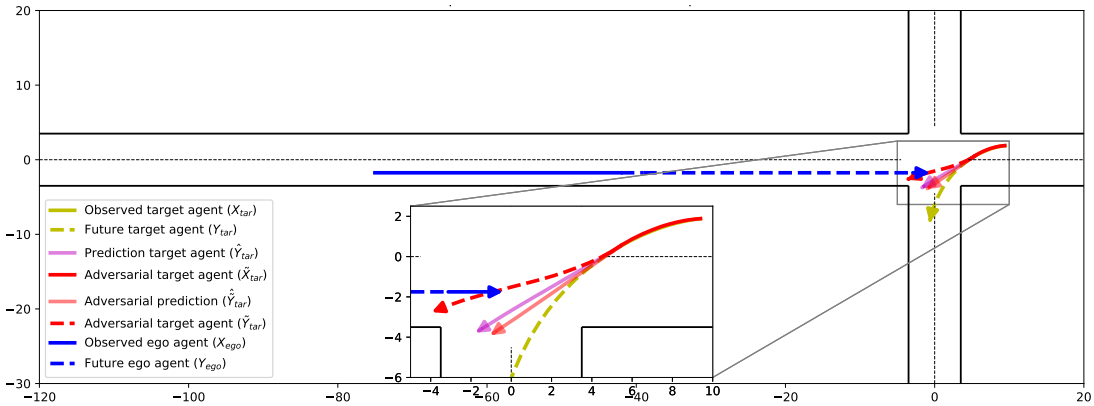

Fig. 19. Perturbing control actions with attack loss $\mathcal{L} = l_{\text{FNC}}(Y_{\text{ego}}, \widetilde{Y}_{\text{tar}}, \widehat{\widetilde{Y}}_{\text{tar}}) + l_{\text{Time}}(\widetilde{X}_{\text{tar}}, X_{\text{tar}})$ The prediction on the adversarial observed states is directed toward the nominal setting prediction. The perturbed future states are directed toward the ego agent, causing a collision.