

Figure 1: Typical case with TTRMARL-GS in the learned environment when  $n_1 = 5$ .

## 1. The learned environment ( $n = 2$ ): Random maneuver

Figure 1 illustrates a typical case showcasing the interceptors' successful interception of the target in the learned environment using TTRMARL-GS. In Figure 1(a), the interceptors  $I_1$  and  $I_2$  intercept the target  $T$  simultaneously, culminating in the interceptor  $I_1$  successfully intercepting the target  $T$  first. In Figure 1(b), the initial relative distance between the interceptor  $I_1$  and the target  $T$  measures approximately 10413 m. Over time, the interceptor  $I_1$  gradually closes the gap, resulting in a steady reduction of the relative distance, eventually converging to approximately 0.425 m. Meanwhile, the interceptor  $I_2$  begins with an initial relative distance of around 10,573 m from the target  $T$ , gradually diminishing it. However, owing to the interceptor  $I_1$ 's successful interception of the target  $T$  beforehand, the interceptor  $I_2$ 's final relative distance settles at approximately 1038.701 m. Furthermore, due to the proximity between the interceptor  $I_1$  and the target  $T$ , a slight jump in LOS angle at the interception point is presented in Figure 1(c). In conclusion, the results underscore the efficacy of TTRMARL-GS.

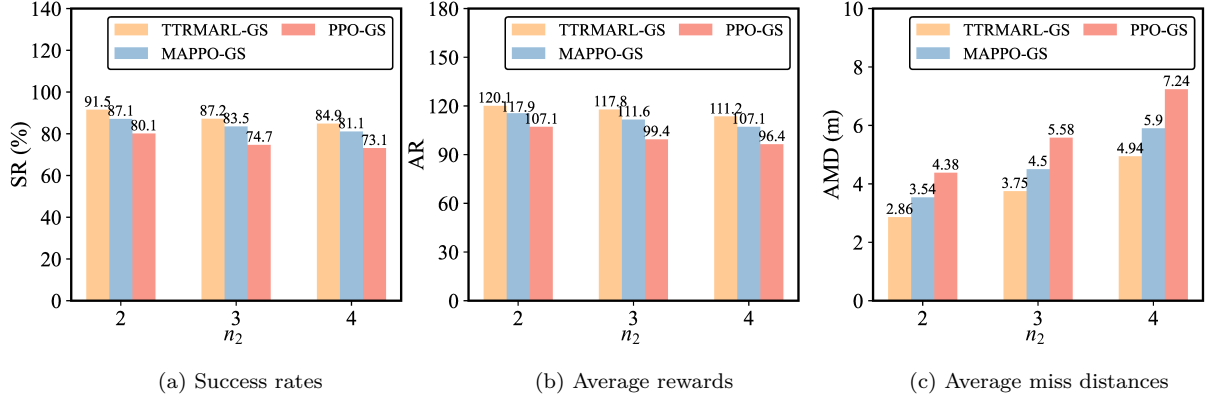


Figure 2: Monte Carlo test results in the unlearned environment I.

## 2. The unlearned environment I ( $n = 2$ ): Step maneuver

Next, TTRMARL-GS, MAPPO-GS and PPO-GS are tested in the unlearned environment I. In the unlearned environment I, the target performs a step maneuver, which is represented as

$$a_T = \begin{cases} 0 & \text{else} \\ n_2 \cdot g & 10 \leq t \leq 16 \end{cases}. \quad (1)$$

The only difference from the learned environment lies in the target performing step maneuver, thereby assessing the interceptors' real-time decision-making ability. All guidance strategies undergo evaluation across 2000 episodes, with the outcomes detailed in Figure 2.

As shown in Figure 2, it is observed that the success rate of all guidance strategies declines with the increase of target maneuver complexity. Nonetheless, noteworthy is that the success rate of TTRMARL-GS remains the highest at 91.5%, 87.2% and 84.9%, outperforming that of MAPPO-GS (87.1%, 83.5% and 81.1%) and PPO-GS (80.1%, 74.7% and 73.1%). Similarly, when  $n_2 = 3$ , TTRMARL-GS exhibits a superior average reward of approximately 117.8 compared to MAPPO-GS (111.6) and PPO-GS (99.4). Furthermore, when  $n_2 = 2, 3, 4$ , the average miss distance of TTRMARL-GS is the closest.

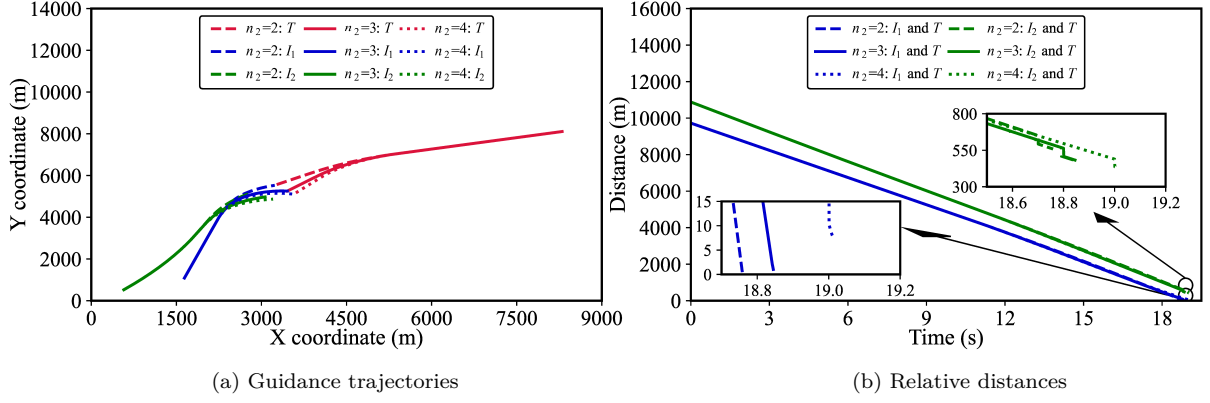


Figure 3: Comparison case with  $n_2 = 2, 3, 4$  in the unlearned environment I.

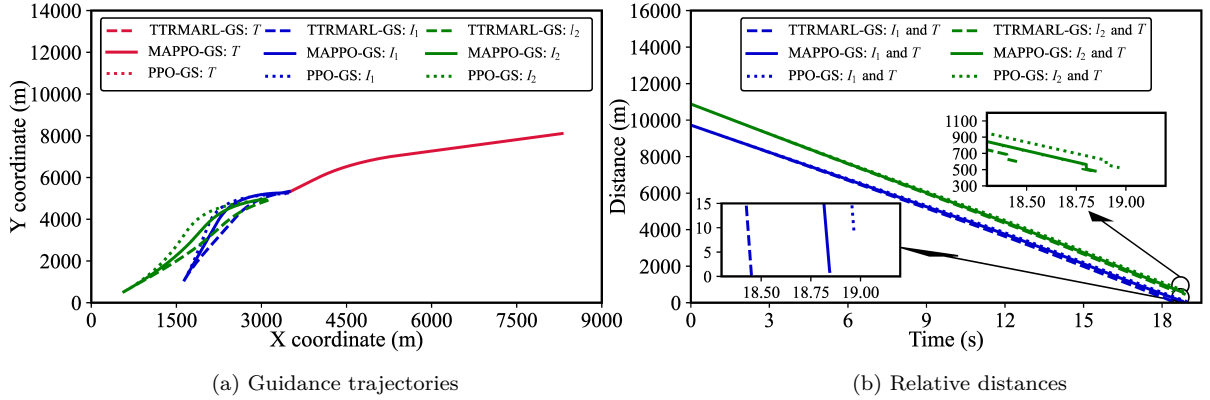


Figure 4: Comparison case with TTRMARL-GS, MAPPO-GS and PPO-GS in the unlearned environment I

These findings underscore the robust generalization ability of TTRMARL-GS.

As depicted in Figure 3 and Figure 4, when  $n_2 = 2, 3, 4$ , the interceptors successfully intercept the target. When  $n_2 = 2$ , the time is minimized, accompanied by the smallest distance. Furthermore, various guidance strategies empower the interceptors to intercept the target effectively, but TTRMARL-GS results in the smallest miss distance. In summary, TTRMARL-GS has remarkable generalization and guidance capabilities.

Figure 5 provides a typical case in which the interceptors employ TTRMARL-GS to intercept the target in the unlearned environment I. In Figure 5(a), the interceptors  $I_1$  and  $I_2$  collaborate in interception, ultimately intercepting the target  $T$ . In Figure 5(b), the initial relative distance between the interceptor  $I_1$  and the target  $T$  measures approximately 9725.799 m. Over time, this relative distance diminishes to 1.033 m.

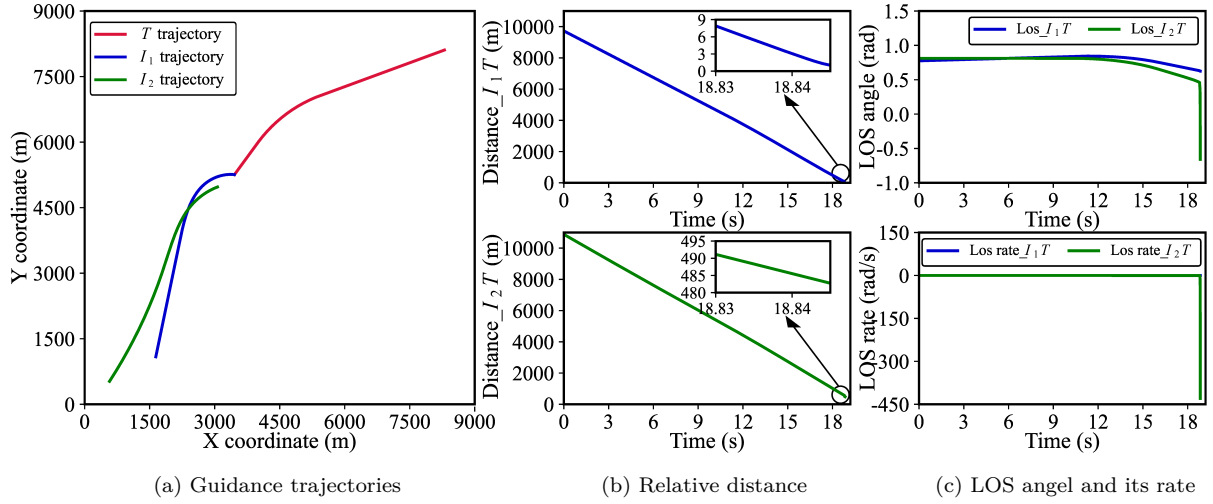


Figure 5: Typical case with TTRMARL-GS in the unlearned environment I when  $n_2 = 3$ .

Conversely, the interceptor  $I_2$  initiates with an initial relative distance of around 7,727 m from the target  $T$ . As the interceptor  $I_1$  successfully intercepts the target  $T$ , the relative distance gradually reduces to 82.733 m, concluding the guidance. In conclusion, these results affirm the efficacy of TTRMARL-GS in effectively intercepting the target.

### 3. The unlearned environment II ( $n = 2$ ): Square wave maneuver

Subsequently, TTRMARL-GS, PPO-GS and MAPPO-GS are subjected to testing in the unlearned environment II, where the target performs a square wave maneuver, which is represented as

$$a_T = n_3 \cdot \text{sign}[\sin(2\pi/T \cdot t)]g. \quad (2)$$

All guidance strategies are evaluated across 2000 episodes, as shown in Figure 6.

Figure 6 shows that the success rate of TTRMARL-GS surpasses that of MAPPO-GS (86.9%, 84.6% and 80.4%) and that of PPO-GS (82.3%, 78.1% and 76.9%). Simultaneously, the average reward achieved by TTRMARL-GS, exceeds that of MAPPO-GS (108.2, 105.6 and 101.9) and PPO-GS (102.6, 100.4 and 94.2). Furthermore, compared to MAPPO-GS and PPO-GS, TTRMARL-GS exhibits the shortest average miss distance.

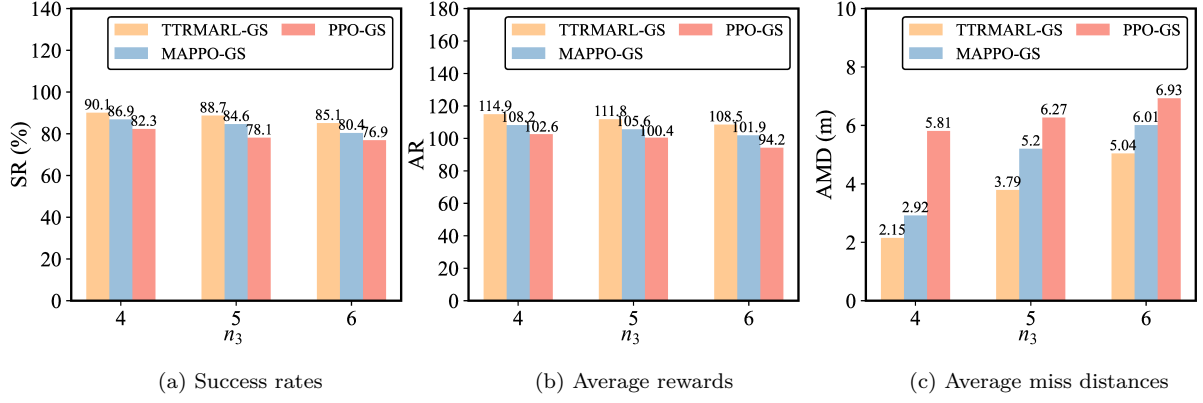


Figure 6: Monte Carlo test results in the unlearned environment II.

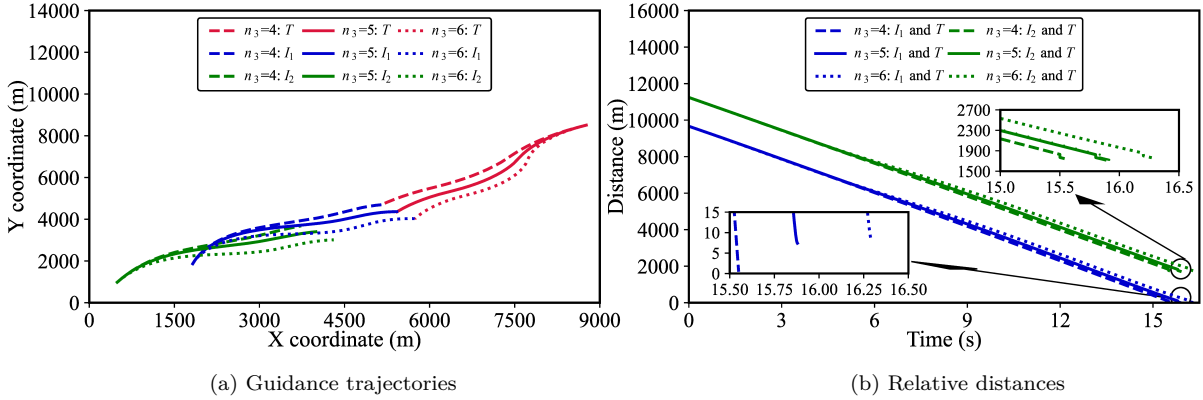


Figure 7: Comparison case with  $n_3 = 4, 5, 6$  in the unlearned environment II.

These results collectively indicate the superior guidance performance of TTRMARL-GS.

Figure 7 illustrates the interception process of various  $n_3$  in the unlearned environment II. It is evident that despite different  $n_3$ , TTRMARL-GS effectively intercepts the target, underscoring the generalization capability of TTRMARL-GS. Figure 8 displays the interception under different guidance strategies in the unlearned environment II. Compared with MAPPO-GS, TTRMARL-GS exhibits a comparatively shorter miss distance, validating its guidance prowess.

Figure 9 presents a typical scenario wherein the interceptors effectively intercept the target employing TTRMARL-GS in the unlearned environment II. Specifically, in Figure 9(a) and Figure 9(b), both interceptors  $I_1$  and  $I_2$  engage simultaneously, while

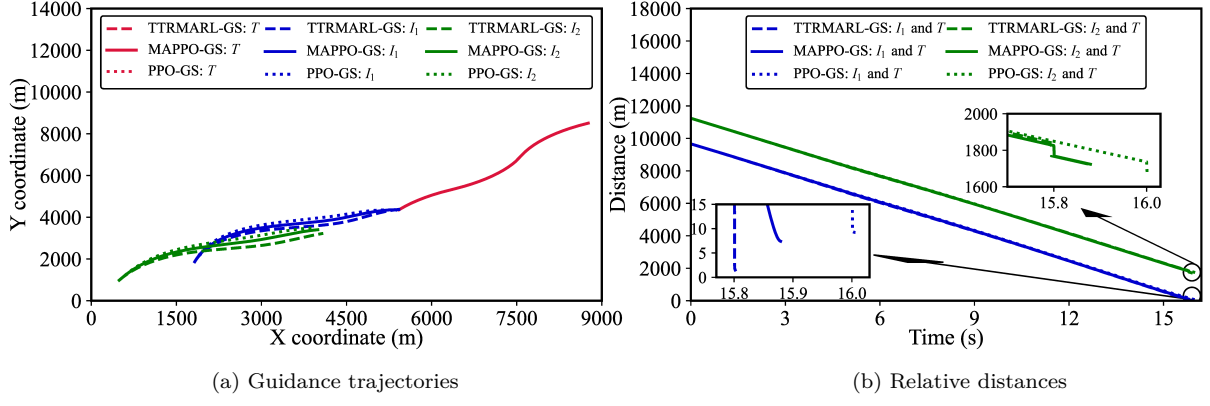


Figure 8: Comparison case with TTRMARL-GS, MAPPO-GS and PPO-GS in the unlearned environment II.

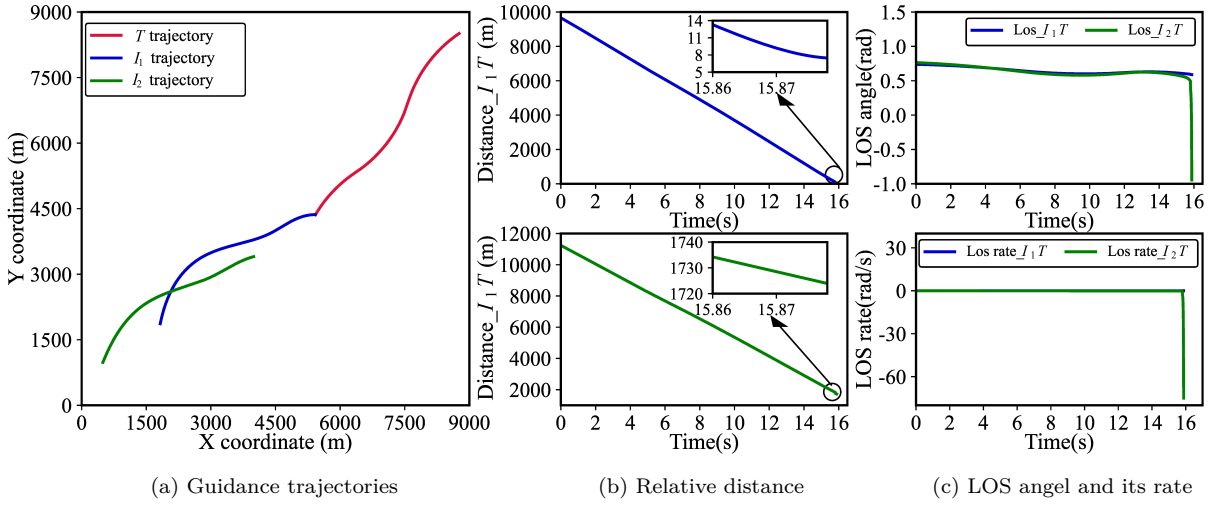


Figure 9: Typical case with TTRMARL-GS in the unlearned environment II when  $n_3 = 5$ .

the interceptor  $I_2$  successfully intercepts the target  $T$ , culminating the guidance. These results highlight the efficacy of TTRMARL-GS in successfully intercepting the target.