Table 1: The FPS of PolyDiffuse counts both the MapTR proposal generator and the GS-DM. The running time in this table is measured on a single Nvidia RTX A5000 GPU. The PolyDiffuse here is retrained during rebuttal and has a slightly better performance than Table.2 of the main paper.

| Matching Criterion → | | | | Chamfer distance | | | | + Ordered angle distance | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Stages | Steps | FPS | $AP_p$ | $AP_d$ | $AP_b$ | mAP | $AP_p$ | $AP_d$ | $AP_b$ | mAP |
| MapTR | 1 | 1 | 14.3 | 55.8 | **60.9** | 61.1 | 59.3 | 46.1 | 43.4 | 41.9 | 43.8 |
| +PolyDiffuse | 2 | 2 | 6.3 | 56.8 | 59.8 | 60.9 | 59.2 | 50.3 | 48.2 | 44.3 | 47.6 |
| +PolyDiffuse | 2 | 5 | 4.8 | 58.1 | 59.7 | 61.2 | 59.6 | 51.8 | 49.5 | 45.4 | 48.9 |
| +PolyDiffuse | 2 | 10 | 3.4 | **58.2** | 59.7 | **61.3** | **59.7** | **52.0** | **49.5** | **45.4** | **49.0** |

Table 2: The FPS of PolyDiffuse counts both the RoomFormer proposal generator and the GS-DM. The running time is measured with a single Nvidia RTX A5000 GPU. The same PolyDiffuse model is used as the Table.1 of the main paper.

| Evaluation Level → | | | | Room | | | Corner | | | Angle | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Method | Stages | Steps | FPS | Prec. | Rec. | F1 | Prec. | Rec. | F1 | Prec. | Rec. | F1 |
| RoomFormer | 1 | 1 | 29.9 | 96.3 | 96.2 | 96.2 | 89.7 | 86.7 | 88.2 | 85.4 | 82.5 | 83.9 |
| +PolyDiffuse | 2 | 2 | 11.7 | 96.9 | 96.4 | 96.6 | 90.3 | 87.1 | 88.7 | 85.8 | 82.8 | 84.3 |
| +PolyDiffuse | 2 | 5 | 7.1 | 98.5 | 97.9 | 98.2 | 92.5 | 89.0 | 90.7 | 90.3 | 86.9 | 88.6 |
| +PolyDiffuse | 2 | 10 | 4.4 | **98.7** | **98.1** | **98.4** | **92.8** | **89.3** | **91.0** | **90.8** | **87.4** | **89.1** |

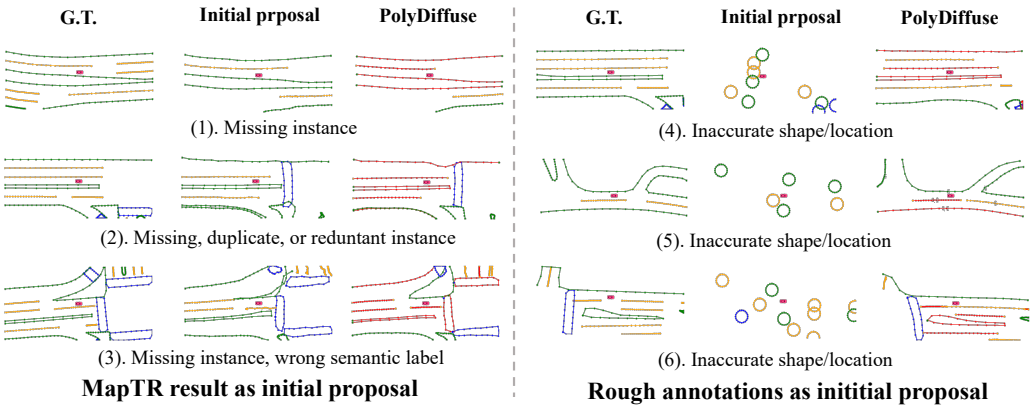


Figure 1: Failure examples of PolyDiffuse on the HD mapping task. We mark a predicted map element by red color if it is a true positive under both the Chamfer distance and the order-aware angle distance matching criteria. The thresholds used here are 1.0m and 10°, respectively. See the text in the global response for discussions and analyses.

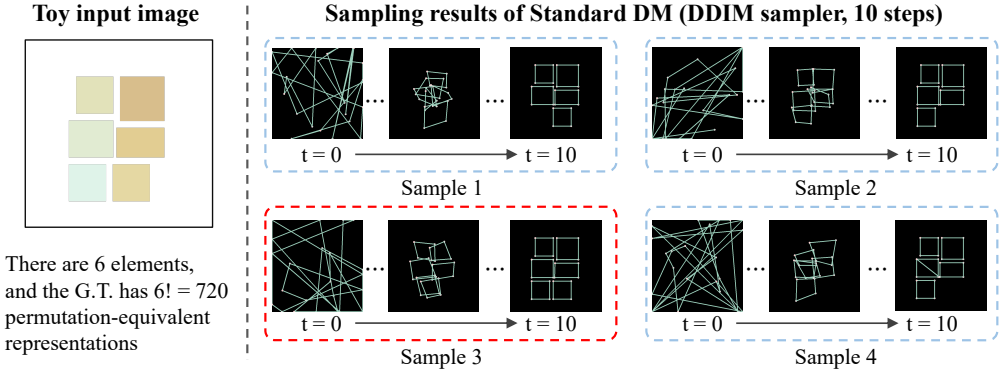


Figure 2: A simple toy experiment of using a standard DM to fit a single data sample with 6 elements. Four sampling results with *different initial noises* are shown. The DDIM sampler is used with 10 denoising steps. Only one of the four samples (*i.e.*, Sample 3) gets the correct final result due to the challenges induced by the set ambiguity, as explained in the main paper.