SIMSCALE: Learning to Drive via Real-World Simulation at Scale

Supplementary Material

We outline the supplementary material as follows. In Sec. A, we first provide additional discussions of related work. Sec. B presents further implementation details regarding data curation and models. In Sec. C, we report **extended experimental analyses and qualitative results** corresponding to the main experiments (Sec. 3). In Sec. D, we further present **additional experimental studies**. Finally, Sec. E discusses the limitations and broader impacts, and Sec. F lists the licenses of all utilized assets.

A. Related Work

A.1. End-to-End Autonomous Driving

End-to-end system maps directly from raw sensor inputs to planning [6]. Early works adopt regression-based planning and gradually shift from extra task branches [8, 26, 65] to unifying perception, prediction, and planning under joint supervision [22, 27, 28, 52, 64]. Recent work has moved toward generative approaches. Diffusion-based systems [11, 38, 47, 51, 69] are framed as a conditional denoising process, enabling diverse and high-fidelity trajectories. Concurrently, trajectory scoring has emerged as an efficient alternative, ranking candidate trajectories under spline curves [4, 21], discretized tokens [16, 29], clustered human trajectories [48, 50], or predicted proposals [19, 46, 60, 67]. Moreover, it offers a natural interface for reward- or cost-based optimization [16–18, 32, 47, 49, 53], enabling reinforced improvements. Still, supervision on logged end-to-end data limits training to the expert's open-loop distribution, causing compromised learning in the drifted state of sensory data. We address this by introducing a scalable simulation framework that reactively generates end-to-end pairs using pseudo-experts or rewards, allowing end-to-end systems across any above paradigms to bootstrap extra supervision from existing training data.

A.2. Scene Simulation for Driving

Scene simulation, including traffic behavior simulation and sensor simulation, has long been a key topic in autonomous driving research. For traffic simulation, existing works [3, 16, 83] leverage rule-based planners like IDM [62], or diffusion-based generators [31, 86] to simulate plausible interaction of traffic agents. For sensor simulation, traditional graphics-based simulators [14, 40] suffer from a significant sim-to-real gap, which limits the real-world deployment of trained planners. Recent data-driven efforts follow two directions: some approaches [1, 44, 63, 70, 75, 84, 85] attempt to generate sensor data in unseen scenarios via video generation models conditioned on 3D bounding boxes, HD maps or BEV maps. Other works focus on scene reconstruction [3, 7, 30, 54, 68, 71, 83] to build photorealistic simulators for novel-view synthesis, using techniques like Neural Radiance Fields (NeRF) and 3D Gaussian Splatting (3DGS). Although these methods achieve impressive visual results, most are primarily designed for closed-loop evaluation or visual augmentation. Our work aims to explore how to use traffic and sensor simulation to generate realistic scenes with feasible demonstrations for planner training, and its impacts on planner performances.

A.3. Data Scaling for Driving

Recent data scaling laws have driven major advances in foundation models [33, 58, 76, 80], but their impact on autonomous driving planning remains underexplored. Prior researches [24, 56, 81] demonstrate that increasing real-world data from thousands to millions of driving logs improves end-to-end planner performance, though improvements diminish. Dense supervision from video predictions improves this data scaling efficiency [45]. For planners operating on abstract BEV representations, industry efforts

[2, 23] demonstrate clear benefits from scaling real-world data and model capacity, while self-play [9, 10, 77] scales reinforcement learning via massive simulation to achieve strong zero-shot sim-to-real transfer. However, most existing approaches rely on costly real-world collection or traffic simulation, which expands only in abstract state spaces rather than raw sensory domains. 3D rasterization provides a lightweight remedy, but suffers from information loss [15]. In contrast, we investigate scaling properties of planning directly through large-scale 3DGS sensory simulation, which bridges abstract traffic simulation and real-world perception, offering a scalable and realistic alternative to real-world data collection.

A.4. Comparison with Online RL

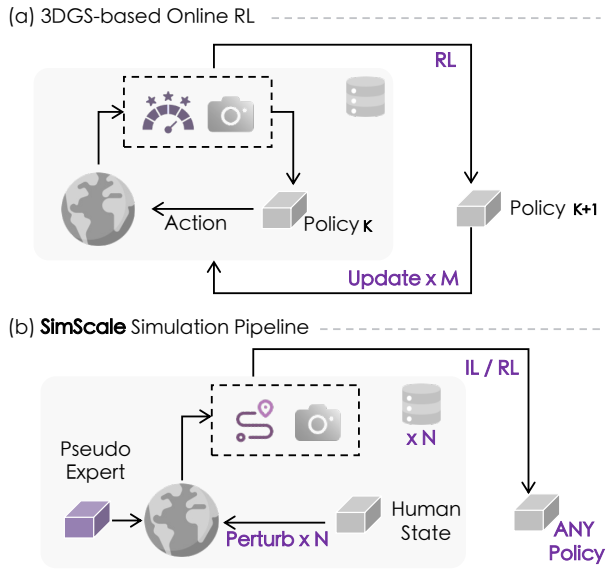


Figure 1. **Learning paradigm comparison of e2e autonomous driving between 3DGS-based Online RL and SimScale.**

Online reinforcement learning (RL) [25, 39, 41] in traffic simulators learn through exploration and feedback. Sensor simulation serves as a bridge to real-world environments, enabling real-world RL for autonomous driving, *i.e.*, 3DGS-based Online RL [16], as shown in Fig. 1 (a). In contrast, the proposed SimScale introduces an alternative framework, as illustrated in Fig. 1 (b), that generates OOD states with expert demonstrations at scale, enabling real-world simulations to support both imitation learning (IL) and reward learning (*i.e.* Offline RL) across arbitrary planners.

B. Extended Implementation Details

B.1. Simulation Data Curation

Tab. 1 summarizes the detailed terms with notation and value used in the simulation data curation process to supplement Sec. 3.1 curation pipeline in the main paper. For pseudo-expert trajectory generation, we discard infeasible candidates to ensure valid supervision. Specifically, all sub-metrics of EPDMS must be satisfied, with S_{EP} relaxed: EP ≥ 0.5, preventing biased driving styles. Qualitative results during curation are shown in Sec. C.2

Table 1. **Simulation Data Curation Pipeline Configurations.**

Term	Notation	Value
3DGS Data Engine		
Peak signal-to-noise ratio of GS	PSNR	< 27
Trajectory Perturbation		
Clustered human traj vocabulary	$\mathcal{V}_{\text{c}}$	$16,384$
Relative heading range to log	$\Delta\theta$	$\pm 20^{\circ}$
Longitudinal shift range	r_{lon}	± 20m
Lateral shift range	r_{lat}	± 2m
Longitudinal sampling step	δ_{lon}	5m
Lateral sampling step	δ_{lat}	0.5m
Reward filterer	$\text{R}_{\text{EPDMS}}^{\text{per}}$	≥ 0.8
Pseudo-Expert Trajectory Generation		
Reward filterer	$\text{R}_{\text{EPDMS}}^{\text{exp}}$	$S_m{=}1, \forall m{\neq}\text{EP}$ $S_{\text{EP}}{>}0.5$
Recovery-based		
Whole human traj vocabulary	$\mathcal{V}_{\text{h}}$	$103,288$
Planner-based		
Privileged planner	**P**	PDM-Closed [12]

Table 2. **Model and Training Hyperparameters.**

Hyperpara.	LTF [8]	DiffusionDrive [51]	GTRS-Dense [50]
Model Configuration			
Sensors	$3 \times$ Cam.	$3 \times$ Cam.	$3 \times$ Cam.
Resolution	2048×512	2048×512	2048×512
Horizon	4s	4s	4s
Frequency	2Hz	2Hz	10Hz
Backbone	R34 [20]	R34 [20]	R34 [20] / V99 [36]
Parameters	56M	61M	67M / 83M
Aux. Tasks	Det. Seg.	Det. Seg.	None
Training Configuration			
GPUs	$8 \times$ H20	$8 \times$ H20	$32 \times$ H20
Epochs	100	100	50
Total BS	512	512	352
Initial LR	2.83×10^{-4}	6×10^{-4}	4×10^{-4}
Schedule	Constant	Cosine Decay	Cosine Decay
Optimizer	Adam [35]	AdamW [55]	AdamW [55]

B.2. Models and Training

Tab. 2 provides the detailed model and training hyperparameters that complement Sec. 3.1 implementation protocol in the main paper. We follow the NAVSIM [13] default setup and use only the `train logs` [34] of `navtrain` for both real-data training and sim–real co-training. All planners share identical settings across the real or sim-real training setups, and each model is trained to saturation to ensure the validity of the subsequent scaling analyses.

LTF. [8] A regression-based planner that directly regresses future waypoints upon fused sensory latents. It employs

Table 3. **Detailed Results on `navhard`.** PDM-Closed uses ground-truth symbolic inputs for planning, while other methods rely on sensor data. (*recovery / planner*: recovery-based / planner-based expert; *reward*: reward-only scoring; S.: per-stage EPDM score.)

Method	Backbone	Sim.	Stage	NC↑	DAC↑	DDC↑	TLC↑	EP↑	TTC↑	LK↑	HC↑	EC↑	S.↑	EPDMS↑
PDM-Closed [12]	-	-	S 1	94.4	98.8	100	99.5	100	93.5	99.3	87.7	36.0	-	51.3
			S 2	88.1	90.6	96.3	98.5	100	83.1	73.7	91.5	25.4	-	
Regeression-based Planner														
LTF [8]	ResNet34	*w/o*	S 1	97.3	80.2	97.8	99.3	83.4	96.2	92.9	97.8	71.1	61.3	24.4
			S 2	79.4	69.0	85.6	98.5	83.8	76.7	47.9	97.0	70.6	39.2	
		recovery	S 1	96.4	78.4	98.9	99.8	80.5	96.2	92.7	97.6	78.2	60.7	29.8
			S 2	88.9	71.1	91.8	99.0	77.3	85.5	53.8	96.8	47.5	**46.3**	
		planner	S 1	96.1	85.3	99.4	99.3	84.7	94.7	93.6	97.6	77.3	**66.3**	**30.2**
			S 2	85.5	66.9	91.6	99.1	93.0	81.1	58.3	95.1	42.9	44.8	
Diffusion-based Planner														
DiffusionDrive [51]	ResNet34	*w/o*	S 1	96.8	86.0	98.8	99.3	84.0	95.8	96.7	97.6	79.6	66.7	27.5
			S 2	80.1	72.8	84.4	98.4	85.9	76.6	46.4	96.3	72.8	40.5	
		recovery	S 1	97.2	88.4	99.1	99.8	83.9	96.0	96.7	97.6	76.9	**69.4**	30.4
			S 2	82.4	67.7	89.1	98.6	89.0	77.6	53.8	95.2	46.8	41.7	
		planner	S 1	97.2	88.7	99.3	99.3	82.8	96.9	98.0	97.3	59.6	67.5	**32.6**
			S 2	82.4	72.1	92.9	98.5	92.1	80.6	60.8	95.4	31.9	**46.8**	
Scoring-based Planner														
GTRS-Dense [50]	ResNet34	*w/o*	S 1	99.3	96.6	99.6	100	57.4	99.5	92.6	89.5	16.4	67.1	38.3
			S 2	92.8	88.6	95.5	99.4	55.9	91.3	55.7	91.1	35.7	55.8	
		recovery	S 1	97.6	94.2	99.3	99.6	76.7	97.6	94.9	97.1	37.8	69.8	43.0
			S 2	92.0	88.1	94.3	98.6	83.6	89.8	58.1	91.4	25.9	59.3	
		planner	S 1	98.7	94.2	99.7	100	74.3	98.2	95.8	96.7	34.7	70.8	46.1
			S 2	94.2	92.1	95.0	99.0	79.9	91.8	58.9	91.0	32.9	**63.9**	
		reward	S 1	97.6	96.4	99.3	100	75.7	97.8	93.3	97.3	32.9	**72.4**	**46.9**
			S 2	94.3	92.7	95.1	99.5	80.2	91.5	56.2	90.6	28.3	63.4	
	V2-99	*w/o*	S 1	98.9	94.9	99.1	100	76.1	98.4	93.8	94.9	37.8	70.4	41.9
			S 2	89.9	90.5	94.1	99.3	77.6	88.5	56.0	92.0	30.2	58.5	
		recovery	S 1	99.1	98.2	99.4	100	71.9	99.1	95.6	95.8	32.4	**73.4**	46.4
			S 2	95.0	90.8	94.6	99.4	75.5	93.5	57.4	93.7	36.8	63.1	
		planner	S 1	99.3	97.1	99.9	100	67.2	99.3	94.0	94.4	23.6	71.7	47.7
			S 2	95.6	91.9	95.0	98.9	76.7	93.7	61.9	90.8	38.0	**65.5**	
		reward	S 1	99.3	97.6	99.4	100	71.5	99.6	96.0	95.8	30.7	72.5	**48.0**
			S 2	94.9	94.3	94.5	99.3	79.0	92.6	57.8	93.4	30.9	65.4	

Table 4. **Detailed Results on `navtest`.** (*recovery / planner*: recovery-based / planner-based expert; *reward*: reward-only scoring)

Method	Backbone	Sim.	NC↑	DAC↑	DDC↑	TLC↑	EP↑	TTC↑	LK↑	HC↑	EC↑	EPDMS↑
Human Agent	-	-	100	100	99.8	100	87.4	100	100	98.1	90.1	90.3
Regression-based Planner												
LTF [8]	ResNet34	*w/o*	97.7	94.0	99.3	99.8	87.2	96.7	95.5	98.3	82.9	81.5
		recovery	97.9	95.1	99.5	99.8	87.7	97.2	97.2	98.3	87.1	83.6
		planner	98.3	95.6	99.6	99.8	87.1	97.5	97.2	98.3	88.2	**84.4**
Diffusion-based Planner												
DiffusionDrive [51]	ResNet34	*w/o*	98.4	95.5	99.5	99.8	87.5	97.5	96.9	98.4	87.7	84.2
		recovery	98.4	96.7	99.6	99.8	87.6	97.5	97.5	98.3	87.1	85.4
		planner	98.5	97.1	99.6	99.8	87.4	97.8	98.0	98.3	87.5	**85.9**
Scoring-based Planner												
GTRS-Dense [50]	ResNet34	*w/o*	97.6	97.5	99.0	99.9	87.9	97.0	95.9	97.5	55.9	82.3
		recovery	98.2	97.9	99.5	99.9	87.2	97.8	96.3	97.6	56.5	83.4
		planner	97.8	98.3	99.5	99.9	89.2	97.3	97.3	96.9	55.1	84.0
		reward	98.4	98.8	99.4	99.9	88.0	98.0	96.5	97.6	58.0	**84.6**
	V2-99	*w/o*	97.6	98.5	99.5	99.9	89.5	97.2	96.8	97.2	57.2	84.0
		recovery	98.3	98.7	99.4	99.9	88.9	97.9	97.1	97.6	55.9	84.5
		planner	98.7	99.0	99.5	99.8	86.8	98.3	95.9	97.9	58.5	84.5
		reward	97.6	98.9	99.6	99.9	89.9	97.1	97.2	97.7	58.7	**84.8**

pretrained ResNet34 [20] as image encoder and is trained for 100 epochs on 8 GPUs with a total batch size of 512 and a learning rate of 2.83×10^{-4}.

DiffusionDrive. [51] A multi-modal, DETR-style generative planner that iteratively denoises diverse trajectories using anchor-conditioned truncated diffusion. Each anchor adaptively queries the encoded image features. It employs pretrained ResNet34 [20] as image encoder and applies truncated diffusion with 20 clustered anchors. It is trained for 100 epochs on 8 GPUs with a total batch size of 512 and a learning rate of 6×10^{-4}.

GTRS-Dense. [50] A scoring-based, DETR-style planner that ranks a dense vocabulary of clustered human trajectories using multiple supervised scoring heads. Inputs are augmented with random spatial shifts. It employs pretrained ResNet34 [20] or V2-99 [36] as image encoder with an action space of 16,384 trajectories. It is trained for 50 epochs on 32 GPUs with a total batch size of 352 and a learning rate of 4×10^{-4}. We use a vocabulary of 16,384 trajectories for training. At inference time, 8,192 trajectories are used for `navhard`, while the full 16,384 trajectories are used for `navtest`.

C. Extended Experimental Results

C.1. Detailed Leaderboard Results

In the main paper Sec. 3.2, Tab. 1 and Tab. 2 report only the best results in `navhard` and `navtest` of each planner co-training (highlighted in blue rows). Supplementary Tab. 3 and Tab. 4 provide full results across different pseudo-experts and the reward-only scoring method. The best results remain highlighted in blue rows, allowing direct comparison with the main paper. Planner-based experts generally outperform recovery-based experts, while scoring-based planner benefits most from reward-only method.

C.2. Complete Qualitative Results

We show extended visualizations of our simulated data in Fig. 5 and Figs. 6–8, which demonstrate the high fidelity and pseudo-expert generation of our data generation pipeline. More OOD simulation scenes are shown in Fig 9.

Recovery-based Expert. As shown in Fig. 5, all simulation trajectories successfully converge back to the human expert's final waypoint.

Planner-based Expert As shown in Figs. 6–8, the simulated feasible trajectories under perturbed states exhibit substantial variation and broader coverage across diverse traffic scenarios, including the Y-branch intersection, the dense urban avenue, and the narrow local road.

C.3. Scaling of Reactive vs. Non-Reactive

In the main paper, Tab. 4 reports only a subset of simulation rounds for GTRS-Dense. Supplementary Fig. 2 provides the full scaling curves, demonstrating that reactive simulation data consistently achieves superior scaling performance across model sizes.

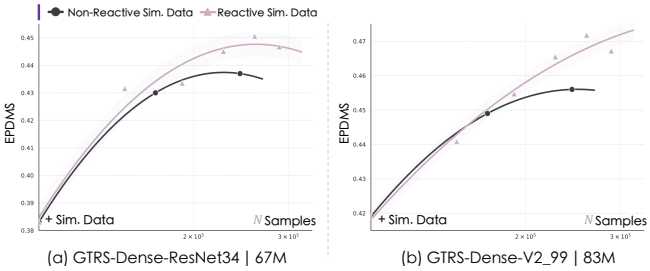


Figure 2. **Scaling dynamics under reactive and non-reactive simulation** using GTRS-Dense across model sizes.

D. Additional Experimental Results

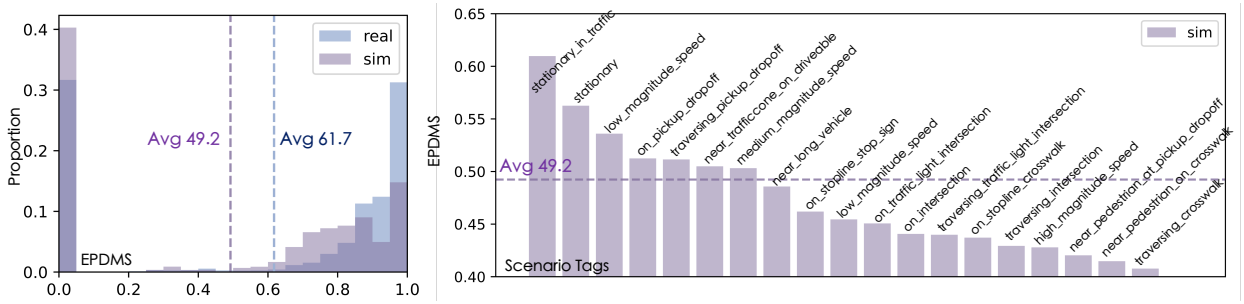
D.1. Multi-Expert Ensemble

Table 5. **The effect of multi-expert ensemble** on `navhard` using GTRS-Dense. ($\mathbf{S}_{1/2}$:Per-stage EPDM scores; *recovery / planner*: recovery-based / planner-based expert; *reward*: reward scoring only; $*$:multi-expert ensemble)

Backbone	Sim	Expert	$\mathbf{S}_1 \uparrow$	$\mathbf{S}_2 \uparrow$	EPDMS $\uparrow$
ResNet34			67.1	55.8	38.3
	✓	*recovery*	69.8	59.3	43.0
	✓	*planner*	70.8	63.9	46.1
	✓	*reward*	72.4	63.4	46.9
	✓	$*$	**72.8**	**65.1**	**47.7**
V2-99			70.4	58.5	41.9
	✓	*recovery*	73.4	63.1	46.4
	✓	*planner*	71.7	65.5	47.7
	✓	*reward*	72.5	65.4	48.0
	✓	$*$	**74.8**	**67.6**	**50.9**
$*$	✓	$*$	75.4	66.4	50.4

Different pseudo-experts exhibit distinct behavioral characteristics, *e.g.*, the recovery-based expert is more conservative, whereas the planner-based expert is more exploratory. This raises the question of whether these complementary properties can be leveraged jointly. Therefore, we conduct a simple ensemble study, multi-expert enseble as we call it, to examine their potential complementarity on `navhard`, as shown in Tab. 5. The scoring-based GTRS-Dense planner is chosen because it enables a straightforward ensemble: the predicted sub-scores of each trajectory in the vocabulary can be directly averaged across the three models (*recovery-based*, *planner-based*, *reward-only scoring*).

Ensemble Gains from Expert Diversity. As shown in Tab. 5, this simple multi-expert ensemble yields consistent



(a) EPDMS Distribution of EgoMLP on Sim. and Real Data. (b) EPDMS of EgoMLP over the Top-20 Most Frequent Scenarios in Sim Data

Figure 3. **EgoMLP** EPDMS **on simulation and real data.** (a) EPDMS distribution under simulation versus real-world data. (b) Scenario tags ranked by EPDMS. The figures reveal the distribution shift and scenario characteristics of simulation data.

improvements for both backbones, achieving +0.8 and +2.9 EPDMS on ResNet34 and V2-99 compared with reward-only scoring, respectively, for `navhard`. Further enlarging the ensemble to all six models yields no additional EPDMS gains. These indicate that, for end-to-end planning, different pseudo-experts provide strong complementary benefits—often exceeding the gains brought by different backbone structures.

D.2. Effect of Scenarios beyond Scaling

Beyond data scaling, simulation data exhibits an out-of-distribution shift relative to real-world data. We provide a preliminary analysis of these characteristics by evaluating an EgoMLP planner, trained on `navtrain`, on both `navtest` and simulation data, as shown in Fig. 3. (1) Fig. 3 (a): Sim data shows a lower EPDMS distribution than real data, confirming a higher safety-critical level. (2) Fig. 3 (b): Low EPDMS scores across nuPlan scenario tags identify challenging simulation scenarios, facilitating future study on their distinct impact, *e.g.*, increasing their proportions during simulation.

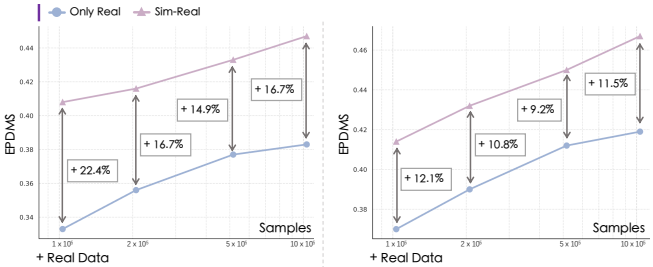
D.3. Effect of Simulation Visual Fidelity

Table 6. **Effect of simulation visual fidelity** on `navhard` using GTRS-Dense, across sampling rounds and co-training strategies.

Backbone	PSNR	**Pseudo-Expert**		**Reward-Only Scoring**	
		#R 1	**#R 2**	**#R 1**	**#R 2**
ResNet34	< 27	41.0	41.1	40.8	41.6
	≥ 27	**41.9**	**41.2**	**41.3**	**42.9**
V2-99	< 27	44.4	44.5	43.7	44.8
	≥ 27	**45.4**	**46.1**	**44.9**	**46.3**

We retrain GTRS-Dense with new simulation data of PSNR < 27 ($\sim$10K$\times$2 rounds) and compare it with the matched samples randomly selected from simulations with PSNR ≥ 27. The former represents lower reconstruction quality and is more likely to yield simulations with reduced visual fidelity. As shown in Tab. 6, simulation data with PSNR ≥ 27 consistently yields higher EPDMS under different data scales and co-training strategies. These results indicate that simulation data with higher visual fidelity helps reduce the visual gap between simulation and real-world observations, thereby improving the effectiveness of simulation data.

D.4. Scaling with Varying Scales of Real Data



(a) GTRS-Dense-ResNet34 | 67M (b) GTRS-Dense-V2_99 | 83M

Figure 4. **Scaling simulation with varying real data.** Simulation data are scaled by corresponding real data scenario tokens and fixed sim-real data ratio

To assess the practical utility of our approach, we study how simulation data affects performance as the amount of available real-world data scales up. Based on GTRS-Dense with reward-only scoring sim-real co-training, we conduct experiments with different real data from `navtrain` (10K, 20K, 50K, 100K). Inspired by [1], simulation data are generated from identical scenarios of corresponding real data tokens, and we fix the sim-real data ratio (five sampling rounds) for each experiment.

Sustained Simulation Gains Across Real Data Scales. As shown in Fig. 4, our simulation data provides the most significant gains when real data is scarce (10K), achieving +22.4% EPDMS on ResNet34 and +12.1% EPDMS on V2-99 for `navhard`. As the amount of real data increases (from 10K to 100K), these gains remain consistently high, without noticeable narrowing. This demonstrates that our simulation data can effectively complement limited real-

world data and still offers substantial potential to further exploit and amplify the value of real data even when it is abundant.

E. Limitations and Broader Impact

E.1. Pseudo-Expert

Despite its effectiveness, our pseudo-expert generation pipeline has limitations. Current rule-based trajectory perturbations are static. A potential enhancement is the self-evolving approach [74], using pretrained planners to iteratively explore the simulator and collect recovery data. Additionally, the privileged BEV planner we use is rule-based with limited performance, causing some degradation in comfort metrics (HC, EC in Tab. 3) and failing in extremely corner cases. More advanced learning-based BEV planners [25, 61, 82] could further improve generation efficiency and realism.

E.2. Scene Simulation

For traffic behavior simulation in our decoupled scene simulation paradigm, other agents are controlled by IDM [62], which enables interaction but limits scenario diversity. Diffusion-based traffic generators [31, 42, 86] offer a promising improvement. For sensor simulation, feedforward GS [5, 72] can improve reconstruction efficiency, while adding modalities like LiDAR [37] provides complementary modality information. Furthermore, unified world models [59, 66, 73, 78, 79] can serve as a substitute for both behavior and sensor simulation.

E.3. Self-Play

Self-play [9, 10] allows the ego and surrounding agents to share learned policies instead of relying on separate traffic behavior simulators, enabling behaviors to co-evolve through interaction. With sensor simulation, it can support end-to-end policy learning and potentially improve robustness in long-tail scenarios, though it remains limited by interaction cost and simulation fidelity.

E.4. Societal Impact

Despite promising improvements, pseudo-expert scene simulation still has room for quality and efficiency enhancements for real-world deployment. Co-training may be affected by unrealistic simulation visuals and distributional differences in real-world corner cases, which could lead to potential risks. We hope SimScale will inspire both academia and industry in driving and robotics to leverage real-world simulation to address rare corner cases in data scaling, advancing fully autonomous systems that are robust and generalizable. In addition, we will provide the community with a fully open-source simulation dataset and training framework to promote academic research on autonomous driving in simulation.

F. License of Assets

All training and evaluation are performed on data from publicly licensed datasets [3, 34, 57]. The real-world data engine is based on the MTGS codebase [43] under the Apache-2.0 license. We adopt publicly available end-to-end planners, including GTRS [50] (Apache-2.0) and DiffusionDrive [51] and LTF [8] (MIT).

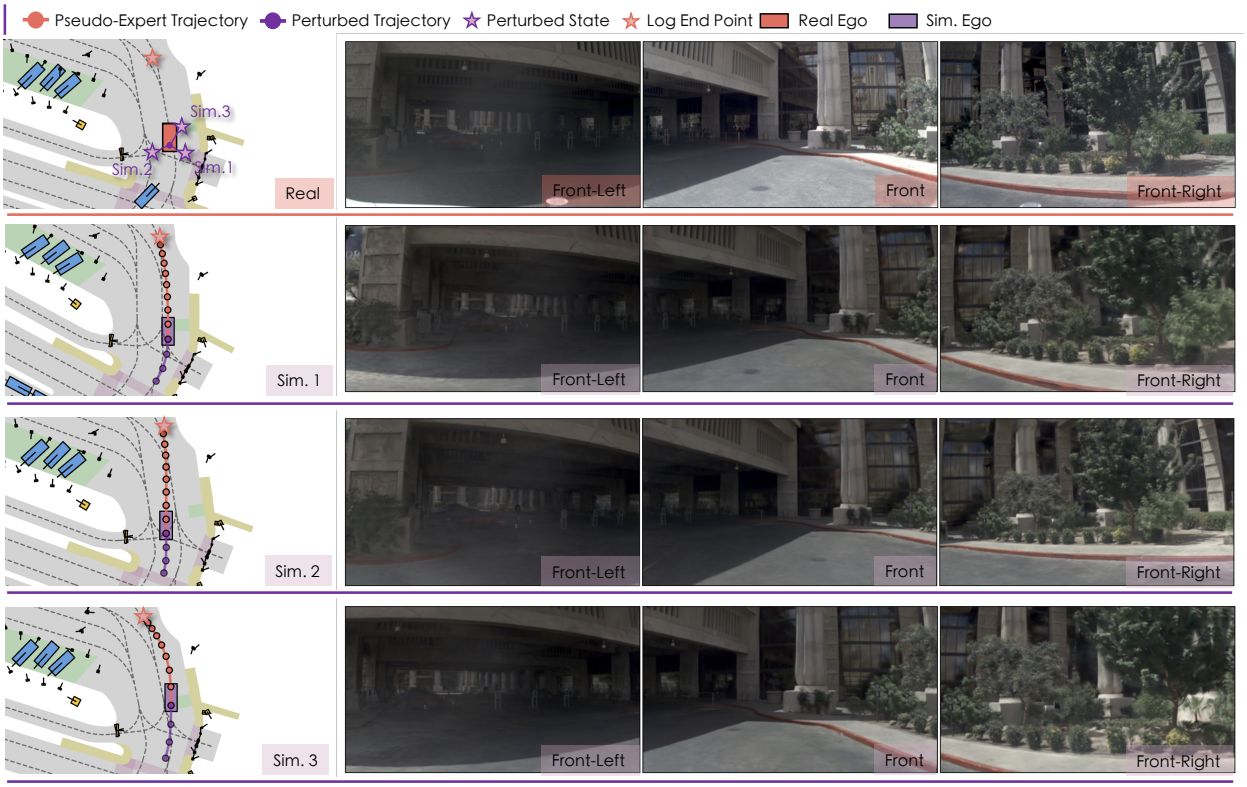


Figure 5. **Qualitative results of recovery-based expert with real and simulation data.**

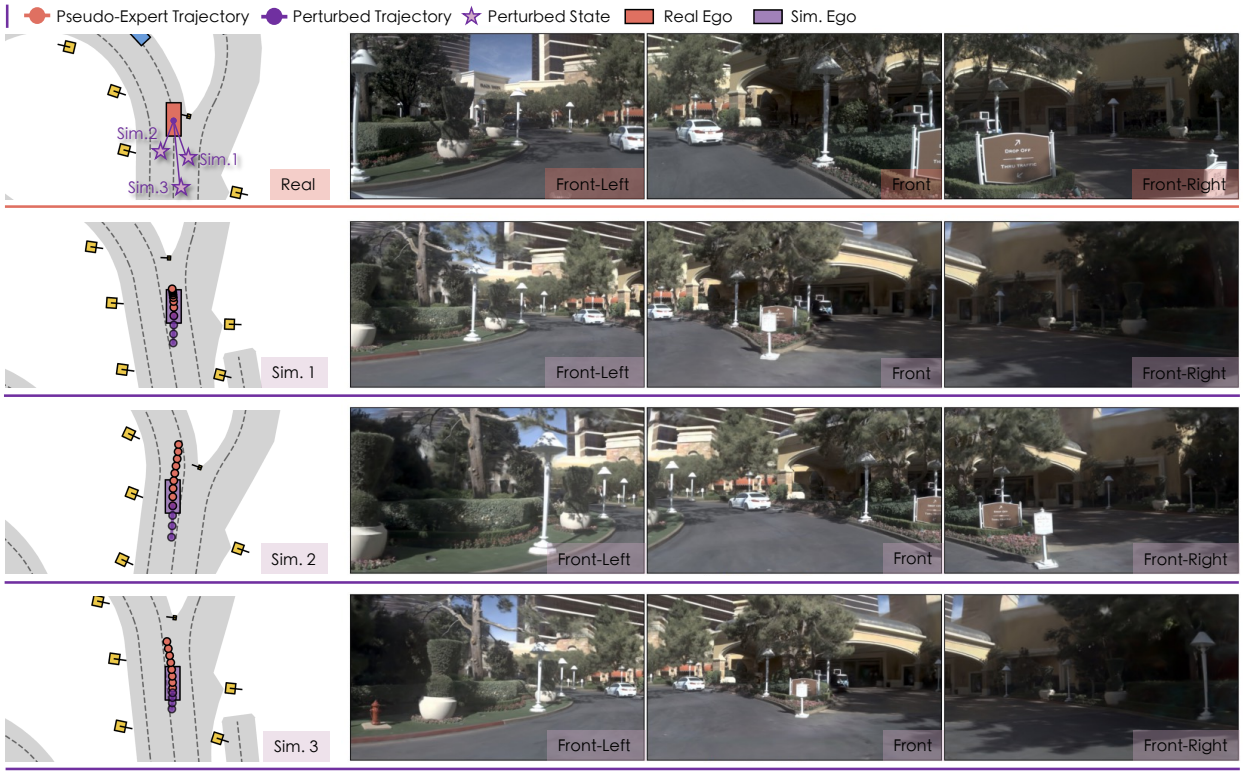


Figure 6. **Qualitative results of planner-based expert with real and simulation data. (Y-branch intersection)**

Figure 7. **Qualitative results of planner-based expert with real and simulation data. (dense urban avenue)**

Figure 8. **Qualitative results of planner-based expert with real and simulation data. (narrow local road)**

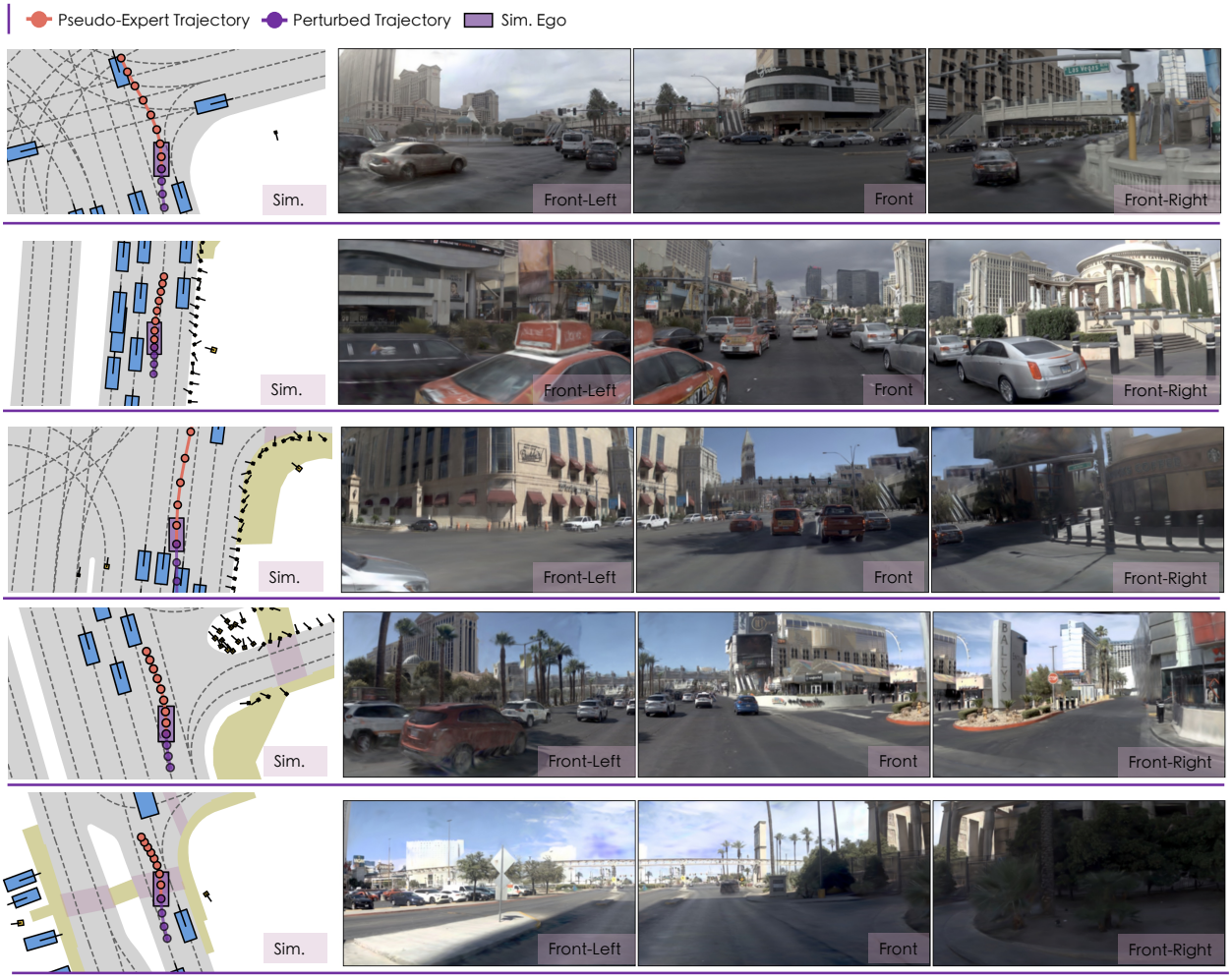


Figure 9. **Additional qualitative results of the simulation scenes on** `navtrain`.

References

[1] Hassan Abu Alhaija, Jose Alvarez, Maciej Bala, Tiffany Cai, Tianshi Cao, Liz Cha, Joshua Chen, Mike Chen, Francesco Ferroni, Sanja Fidler, et al. Cosmos-transfer1: Conditional world generation with adaptive multimodal control. *arXiv preprint arXiv:2503.14492*, 2025. 1, 5

[2] Mustafa Baniodeh, Kratarth Goel, Scott Ettinger, Carlos Fuertes, Ari Seff, Tim Shen, Cole Gulino, Chenjie Yang, Ghassen Jerfel, Dokook Choe, et al. Scaling laws of motion forecasting and planning–a technical report. *arXiv preprint arXiv:2506.08228*, 2025. 2

[3] Wei Cao, Marcel Hallgarten, Tianyu Li, Daniel Dauner, Xunjiang Gu, Caojun Wang, Yakov Miron, Marco Aiello, Hongyang Li, Igor Gilitschenski, Boris Ivanovic, Marco Pavone, Andreas Geiger, and Kashyap Chitta. Pseudo-simulation for autonomous driving. In *CoRL*, 2025. 1, 6

[4] Sergio Casas, Abbas Sadat, and Raquel Urtasun. Mp3: A unified model to map, perceive, predict and plan. In *CVPR*, 2021. 1

[5] David Charatan, Sizhe Li, Andrea Tagliasacchi, and Vincent Sitzmann. pixelsplat: 3d gaussian splats from image pairs for scalable generalizable 3d reconstruction. In *CVPR*, 2024. 6

[6] Li Chen, Penghao Wu, Kashyap Chitta, Bernhard Jaeger, Andreas Geiger, and Hongyang Li. End-to-end autonomous driving: Challenges and frontiers. *TPAMI*, 2024. 1

[7] Ziyu Chen, Jiawei Yang, Jiahui Huang, Riccardo de Lutio, Janick Martinez Esturo, Boris Ivanovic, Or Litany, Zan Gojcic, Sanja Fidler, Marco Pavone, Li Song, and Yue Wang. Omnire: Omni urban scene reconstruction. In *ICLR*, 2025. 1

[8] Kashyap Chitta, Aditya Prakash, Bernhard Jaeger, Zehao Yu, Katrin Renz, and Andreas Geiger. Transfuser: Imitation with transformer-based sensor fusion for autonomous driving. *TPAMI*, 2022. 1, 2, 3, 6

[9] Daphne Cornelisse, Aarav Pandya, Kevin Joseph, Joseph Suárez, and Eugene Vinitsky. Building reliable sim driving agents by scaling self-play. *arXiv preprint arXiv:2502.14706*, 2025. 2, 6

[10] Marco Cusumano-Towner, David Hafner, Alex Hertzberg, Brody Huval, Aleksei Petrenko, Eugene Vinitsky, Erik Wijmans, Taylor Killian, Stuart Bowers, Ozan Sener, et al. Robust autonomy emerges from self-play. In *ICML*, 2025. 2, 6

[11] Chenxu Dang, Sining Ang, Yongkang Li, Haochen Tian, Jie Wang, Guang Li, Hangjun Ye, Jie Ma, Long Chen, and Yan Wang. Drivefine: Refining-augmented masked diffusion vla for precise and robust driving. *arXiv preprint arXiv:2602.14577*, 2026. 1

[12] Daniel Dauner, Marcel Hallgarten, Andreas Geiger, and Kashyap Chitta. Parting with misconceptions about learning-based vehicle motion planning. In *CoRL*, 2023. 2, 3

[13] Daniel Dauner, Marcel Hallgarten, Tianyu Li, Xinshuo Weng, Zhiyu Huang, Zetong Yang, Hongyang Li, Igor Gilitschenski, Boris Ivanovic, Marco Pavone, Andreas Geiger, and Kashyap Chitta. Navsim: Data-driven non-reactive autonomous vehicle simulation and benchmarking. In *NeurIPS*, 2024. 2

[14] Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. CARLA: An open urban driving simulator. In *CoRL*, 2017. 1

[15] Lan Feng, Yang Gao, Eloi Zablocki, Quanyi Li, Wuyang Li, Sichao Liu, Matthieu Cord, and Alexandre Alahi. Rap: 3d rasterization augmented end-to-end planning. *arXiv preprint arXiv:2510.04333*, 2025. 2

[16] Hao Gao, Shaoyu Chen, Bo Jiang, Bencheng Liao, Yiang Shi, Xiaoyang Guo, Yuechuan Pu, Haoran Yin, Xiangyu Li, Xinbang Zhang, Ying Zhang, Wenyu Liu, Qian Zhang, and Xinggang Wang. Rad: Training an end-to-end driving policy via large-scale 3dgs-based reinforcement learning. In *NeurIPS*, 2025. 1, 2

[17] Yinfeng Gao, Deqing Liu, Qichao Zhang, Yupeng Zheng, Haochen Tian, Guang Li, Hangjun Ye, Long Chen, Da-Wei Ding, and Dongbin Zhao. Learning from mistakes: Post-training for driving vla with takeover data. *arXiv preprint arXiv:2603.14972*, 2026.

[18] Yinfeng Gao, Qichao Zhang, Deqing Liu, Zhongpu Xia, Guang Li, Kun Ma, Guang Chen, Hangjun Ye, Long Chen, Da-Wei Ding, and Dongbin Zhao. Perlad: Towards enhanced closed-loop end-to-end autonomous driving with pseudo-simulation-based reinforcement learning. *arXiv preprint arXiv:2603.14908*, 2026. 1

[19] Ke Guo, Haochen Liu, Xiaojun Wu, Jia Pan, and Chen Lv. ipad: Iterative proposal-centric end-to-end autonomous driving. *arXiv preprint arXiv:2505.15111*, 2025. 1

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 2, 4

[21] Shengchao Hu, Li Chen, Penghao Wu, Hongyang Li, Junchi Yan, and Dacheng Tao. St-p3: End-to-end vision-based autonomous driving via spatial-temporal feature learning. In *ECCV*, 2022. 1

[22] Yihan Hu, Jiazhi Yang, Li Chen, Keyu Li, Chonghao Sima, Xizhou Zhu, Siqi Chai, Senyao Du, Tianwei Lin, Wenhai Wang, et al. Planning-oriented autonomous driving. In *CVPR*, 2023. 1

[23] Xin Huang, Eric M Wolff, Paul Vernaza, Tung Phan-Minh, Hongge Chen, David S Hayden, Mark Edmonds, Brian Pierce, Xinxin Chen, Pratik Elias Jacob, et al. Drivegpt: Scaling autoregressive behavior models for driving. *arXiv preprint arXiv:2412.14415*, 2024. 2

[24] Jyh-Jing Hwang, Runsheng Xu, Hubert Lin, Wei-Chih Hung, Jingwei Ji, Kristy Choi, Di Huang, Tong He, Paul Covington, Benjamin Sapp, et al. Emma: End-to-end multimodal model for autonomous driving. *TMLR*, 2025. 1

[25] Bernhard Jaeger, Daniel Dauner, Jens Beißwenger, Simon Gerstenecker, Kashyap Chitta, and Andreas Geiger. Carl: Learning scalable planning policies with simple rewards. In *CoRL*, 2025. 2, 6

[26] Xiaosong Jia, Penghao Wu, Li Chen, Jiangwei Xie, Conghui He, Junchi Yan, and Hongyang Li. Think twice before driving: Towards scalable decoders for end-to-end autonomous driving. In *CVPR*, 2023. 1

[27] Xiaosong Jia, Junqi You, Zhiyuan Zhang, and Junchi Yan. Drivetransformer: Unified transformer for scalable end-to-end autonomous driving. In *ICLR*, 2025. 1

[28] Bo Jiang, Shaoyu Chen, Qing Xu, Bencheng Liao, Jiajie Chen, Helong Zhou, Qian Zhang, Wenyu Liu, Chang Huang, and Xinggang Wang. Vad: Vectorized scene representation for efficient autonomous driving. In *ICCV*, 2023. 1

[29] Bo Jiang, Shaoyu Chen, Hao Gao, Bencheng Liao, Qian Zhang, Wenyu Liu, and Xinggang Wang. VADv2: End-to-end autonomous driving via probabilistic planning. In *ICLR*, 2026. 1

[30] Junzhe Jiang, Nan Song, Jingyu Li, Xiatian Zhu, and Li Zhang. Realengine: Simulating autonomous driving in realistic context. *arXiv preprint arXiv:2505.16902*, 2025. 1

[31] Max Jiang, Yijing Bai, Andre Cornman, Christopher Davis, Xiukun Huang, Hong Jeon, Sakshum Kulshrestha, John Lambert, Shuangyu Li, Xuanyu Zhou, et al. Scenediffuser: Efficient and controllable driving simulation initialization and rollout. In *NeurIPS*, 2024. 1, 6

[32] Siwen Jiao, Kangan Qian, Hao Ye, Yang Zhong, Ziang Luo, Sicong Jiang, Zilin Huang, Yangyi Fang, Jinyu Miao, Zheng Fu, et al. EvaDrive: Evolutionary adversarial policy optimization for end-to-end autonomous driving. *arXiv preprint arXiv:2508.09158*, 2025. 1

[33] Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu, and Dario Amodei. Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*, 2020. 1

[34] Napat Karnchanachari, Dimitris Geromichalos, Kok Seang Tan, Nanxiang Li, Christopher Eriksen, Shakiba Yaghoubi, Noushin Mehdipour, Gianmarco Bernasconi, Whye Kit Fong, Yiluan Guo, et al. Towards learning-based planning: The nuplan benchmark for real-world autonomous driving. In *ICRA*, 2024. 2, 6

[35] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 2

[36] Youngwan Lee and Jongyoul Park. Centermask: Real-time anchor-free instance segmentation. In *CVPR*, 2020. 2, 4

[37] Bohan Li, Jiazhe Guo, Hongsi Liu, Yingshuang Zou, Yikang Ding, Xiwu Chen, Hu Zhu, Feiyang Tan, Chi Zhang, Tiancai Wang, et al. Uniscene: Unified occupancy-centric driving scene generation. In *CVPR*, 2025. 6

[38] Derun Li, Jianwei Ren, Yue Wang, Xin Wen, Pengxiang Li, Leimeng Xu, Kun Zhan, Zhongpu Xia, Peng Jia, Xianpeng Lang, et al. Finetuning generative trajectory model with reinforcement learning from human feedback. *arXiv preprint arXiv:2503.10434*, 2025. 1

[39] Hongchen Li, Tianyu Li, Jiazhi Yang, Haochen Tian, Caojun Wang, Lei Shi, Mingyang Shang, Zengrong Lin, Gaoqiang Wu, Zhihui Hao, et al. Plannerrft: Reinforcing diffusion planners through closed-loop and sample-efficient finetuning. *arXiv preprint arXiv:2601.12901*, 2026. 2

[40] Quanyi Li, Zhenghao Peng, Lan Feng, Qihang Zhang, Zhenghai Xue, and Bolei Zhou. Metadrive: Composing diverse driving scenarios for generalizable reinforcement learning. *TPAMI*, 2022. 1

[41] Qifeng Li, Xiaosong Jia, Shaobo Wang, and Junchi Yan. Think2drive: Efficient reinforcement learning by thinking with latent world model for autonomous driving (in carla-v2). In *ECCV*, 2024. 2

[42] Shihao Li, Naisheng Ye, Tianyu Li, Kashyap Chitta, Tuo An, Peng Su, Boyang Wang, Haiou Liu, Chen Lv, and Hongyang Li. Optimization-guided diffusion for interactive scene generation. *arXiv preprint arXiv:2512.07661*, 2025. 6

[43] Tianyu Li, Yihang Qiu, Zhenhua Wu, Carl Lindström, Peng Su, Matthias Nießner, and Hongyang Li. MTGS: Multi-traversal gaussian splatting. *arXiv preprint arXiv:2503.12552*, 2025. 6

[44] Xiaofan Li, Yifu Zhang, and Xiaoqing Ye. Drivingdiffusion: layout-guided multi-view driving scenarios video generation with latent diffusion model. In *ECCV*, 2024. 1

[45] Yingyan Li, Shuyao Shang, Weisong Liu, Bing Zhan, Haochen Wang, Yuqi Wang, Yuntao Chen, Xiaoman Wang, Yasong An, Chufeng Tang, et al. Drivevla-w0: World models amplify data scaling law in autonomous driving. *arXiv preprint arXiv:2510.12796*, 2025. 1

[46] Yingyan Li, Yuqi Wang, Yang Liu, Jiawei He, Lue Fan, and Zhaoxiang Zhang. End-to-end driving with online trajectory evaluation via bev world model. *arXiv preprint arXiv:2504.01941*, 2025. 1

[47] Yongkang Li, Kaixin Xiong, Xiangyu Guo, Fang Li, Sixu Yan, Gangwei Xu, Lijun Zhou, Long Chen, Haiyang Sun, Bing Wang, et al. ReCogDrive: A reinforced cognitive framework for end-to-end autonomous driving. In *ICLR*, 2026. 1

[48] Zhenxin Li, Shihao Wang, Shiyi Lan, Zhiding Yu, Zuxuan Wu, and Jose M Alvarez. Hydra-next: Robust closed-loop driving with open-loop training. In *ICCV*, 2025. 1

[49] Zhenxin Li, Wenhao Yao, Zi Wang, Xinglong Sun, Jingde Chen, Nadine Chang, Maying Shen, Jingyu Song, Zuxuan Wu, Shiyi Lan, et al. Ztrs: Zero-imitation end-to-end autonomous driving with trajectory scoring. *arXiv preprint arXiv:2510.24108*, 2025. 1

[50] Zhenxin Li, Wenhao Yao, Zi Wang, Xinglong Sun, Joshua Chen, Nadine Chang, Maying Shen, Zuxuan Wu, Shiyi Lan, and Jose M Alvarez. Generalized trajectory scoring for end-to-end multimodal planning. *arXiv preprint arXiv:2506.06664*, 2025. 1, 2, 3, 4, 6

[51] Bencheng Liao, Shaoyu Chen, Haoran Yin, Bo Jiang, Cheng Wang, Sixu Yan, Xinbang Zhang, Xiangyu Li, Ying Zhang, Qian Zhang, et al. DiffusionDrive: Truncated diffusion model for end-to-end autonomous driving. In *CVPR*, 2025. 1, 2, 3, 4, 6

[52] Haochen Liu, Zhiyu Huang, Wenhui Huang, Haohan Yang, Xiaoyu Mo, and Chen Lv. Hybrid-prediction integrated planning for autonomous driving. *TPAMI*, 2025. 1

[53] Haochen Liu, Tianyu Li, Haohan Yang, Li Chen, Caojun Wang, Ke Guo, Haochen Tian, Hongchen Li, Hongyang Li, and Chen Lv. Reinforced refinement with self-aware expansion for end-to-end autonomous driving. *arXiv preprint arXiv:2506.09800*, 2025. 1

[54] William Ljungbergh, Adam Tonderski, Joakim Johnander, Holger Caesar, Kalle Åström, Michael Felsberg, and Christoffer Petersson. Neuroncap: Photorealistic closed-loop safety testing for autonomous driving. In *ECCV*, 2024. 1

[55] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *ICLR*, 2019. 2

[56] Alexander Naumann, Xunjiang Gu, Tolga Dimlioglu, Mariusz Bojarski, Alperen Degirmenci, Alexander Popov, Devansh Bisla, Marco Pavone, Urs Muller, and Boris Ivanovic. Data scaling laws for end-to-end autonomous driving. In *CVPR*, 2025. 1

[57] OpenScene Contributors. OpenScene: The largest up-to-date 3d occupancy prediction benchmark in autonomous driving. https://github.com/OpenDriveLab/OpenScene, 2023. 6

[58] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *ICML*, 2021. 1

[59] Yang Shi, Yuhao Dong, Yue Ding, Yuran Wang, Xuanyu Zhu, Sheng Zhou, Wenting Liu, Haochen Tian, Rundong Wang, Huanqian Wang, et al. Realunify: Do unified models truly benefit from unification? a comprehensive benchmark. *arXiv preprint arXiv:2509.24897*, 2025. 6

[60] Wenchao Sun, Xuewu Lin, Yining Shi, Chuang Zhang, Haoran Wu, and Sifa Zheng. Sparsedrive: End-to-end autonomous driving via sparse scene representation. In *ICRA*, 2025. 1

[61] Tianyi Tan, Yinan Zheng, Ruiming Liang, Zexu Wang, Kexin Zheng, Jinliang Zheng, Jianxiong Li, Xianyuan Zhan, and Jingjing Liu. Flow matching-based autonomous driving planning with advanced interactive behavior modeling. In *NeurIPS*, 2025. 6

[62] Martin Treiber, Ansgar Hennecke, and Dirk Helbing. Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 2000. 1, 6

[63] Xiaofeng Wang, Zheng Zhu, Guan Huang, Xinze Chen, Jiagang Zhu, and Jiwen Lu. Drivedreamer: Towards real-world-drive world models for autonomous driving. In *ECCV*, 2024. 1

[64] Xinshuo Weng, Boris Ivanovic, Yan Wang, Yue Wang, and Marco Pavone. Para-drive: Parallelized architecture for real-time autonomous driving. In *CVPR*, 2024. 1

[65] Penghao Wu, Xiaosong Jia, Li Chen, Junchi Yan, Hongyang Li, and Yu Qiao. Trajectory-guided control prediction for end-to-end autonomous driving: A simple yet strong baseline. In *NeurIPS*, 2022. 1

[66] Yanhao Wu, Haoyang Zhang, Tianwei Lin, Lichao Huang, Shujie Luo, Rui Wu, Congpei Qiu, Wei Ke, and Tong Zhang. Generating multimodal driving scenes via next-scene prediction. In *CVPR*, 2025. 6

[67] Chengen Xie, Bin Sun, Tianyu Li, Junjie Wu, Zhihui Hao, XianPeng Lang, and Hongyang Li. Latentvla: Efficient vision-language models for autonomous driving via latent action prediction. *arXiv preprint arXiv:2601.05611*, 2026. 1

[68] Ziyang Xie, Zhizheng Liu, Zhenghao Peng, Wayne Wu, and Bolei Zhou. Vid2sim: Realistic and interactive simulation from video for urban navigation. In *CVPR*, 2025. 1

[69] Zebin Xing, Xingyu Zhang, Yang Hu, Bo Jiang, Tong He, Qian Zhang, Xiaoxiao Long, and Wei Yin. Goalflow: Goal-driven flow matching for multimodal trajectories generation in end-to-end autonomous driving. In *CVPR*, 2025. 1

[70] Zhiyuan Xu, Bohan Li, Huan-ang Gao, Mingju Gao, Yong Chen, Ming Liu, Chenxu Yan, Hang Zhao, Shuo Feng, and Hao Zhao. Challenger: Affordable adversarial driving video generation. *arXiv preprint arXiv:2505.15880*, 2025. 1

[71] Yunzhi Yan, Haotong Lin, Chenxu Zhou, Weijie Wang, Haiyang Sun, Kun Zhan, Xianpeng Lang, Xiaowei Zhou, and Sida Peng. Street gaussians: Modeling dynamic urban scenes with gaussian splatting. In *ECCV*, 2024. 1

[72] Jiawei Yang, Jiahui Huang, Yuxiao Chen, Yan Wang, Boyi Li, Yurong You, Apoorva Sharma, Maximilian Igl, Peter Karkus, Danfei Xu, et al. Storm: Spatio-temporal reconstruction model for large-scale outdoor scenes. In *ICLR*, 2025. 6

[73] Yan Yang, Haochen Tian, Yang Shi, Wulin Xie, Yi-Fan Zhang, Yuhao Dong, Yibo Hu, Liang Wang, Ran He, Caifeng Shan, et al. A survey of unified multimodal understanding and generation: Advances and challenges. *Authorea Preprints*, 2025. 6

[74] Yifan Ye, Jun Cen, Jing Chen, and Zhihe Lu. Self-evolved imitation learning in simulated world. *arXiv preprint arXiv:2509.19460*, 2025. 6

[75] Junqi You, Xiaosong Jia, Zhiyuan Zhang, Yutao Zhu, and Junchi Yan. Bench2drive-r: Turning real world data into reactive closed-loop autonomous driving benchmark by generative model. *arXiv preprint arXiv:2412.09647*, 2024. 1

[76] Xiaohua Zhai, Alexander Kolesnikov, Neil Houlsby, and Lucas Beyer. Scaling vision transformers. In *CVPR*, 2022. 1

[77] Chris Zhang, Sourav Biswas, Kelvin Wong, Kion Fallah, Lunjun Zhang, Dian Chen, Sergio Casas, and Raquel Urtasun. Learning to drive via asymmetric self-play. In *ECCV*, 2024. 2

[78] Huanyu Zhang, Chengzu Li, Wenshan Wu, Shaoguang Mao, Yifan Zhang, Haochen Tian, Ivan Vulić, Zhang Zhang, Liang Wang, Tieniu Tan, et al. Scaling and beyond: Advancing spatial reasoning in mllms requires new recipes. *arXiv preprint arXiv:2504.15037*, 2025. 6

[79] Huanyu Zhang, Wenshan Wu, Chengzu Li, Ning Shang, Yan Xia, Yangyu Huang, Yifan Zhang, Li Dong, Zhang Zhang, Liang Wang, et al. Latent sketchpad: Sketching visual thoughts to elicit multimodal reasoning in mllms. *arXiv preprint arXiv:2510.24514*, 2025. 6

[80] YiFan Zhang, Huanyu Zhang, Haochen Tian, Chaoyou Fu, Shuangqing Zhang, Junfei Wu, Feng Li, Kun Wang, Qingsong Wen, Zhang Zhang, et al. Mme-realworld: Could your multimodal llm challenge high-resolution real-world scenarios that are difficult for humans? In *ICLR*, 2025. 1

[81] Yupeng Zheng, Zhongpu Xia, Qichao Zhang, Teng Zhang, Ben Lu, Xiaochuang Huo, Chao Han, Yixian Li, Mengjie Yu, Bu Jin, et al. Preliminary investigation into data scaling laws for imitation learning-based end-to-end autonomous driving. *arXiv preprint arXiv:2412.02689*, 2024. 1

[82] Yinan Zheng, Ruiming Liang, Kexin ZHENG, Jinliang Zheng, Liyuan Mao, Jianxiong Li, Weihao Gu, Rui Ai, Shengbo Eben Li, Xianyuan Zhan, and Jingjing Liu. Diffusion-based planning for autonomous driving with flexible guidance. In *ICLR*, 2025. 6

[83] Hongyu Zhou, Longzhong Lin, Jiabao Wang, Yichong Lu, Dongfeng Bai, Bingbing Liu, Yue Wang, Andreas Geiger,

and Yiyi Liao. Hugsim: A real-time, photo-realistic and closed-loop simulator for autonomous driving. *arXiv preprint arXiv:2412.01718*, 2024. 1

[84] Jiawei Zhou, Linye Lyu, Zhuotao Tian, Cheng Zhuo, and Yu Li. Safemvdrive: Multi-view safety-critical driving video synthesis in the real world domain. *arXiv preprint arXiv:2505.17727*, 2025. 1

[85] Yunsong Zhou, Michael Simon, Zhenghao Peng, Sicheng Mo, Hongzi Zhu, Minyi Guo, and Bolei Zhou. Simgen: Simulator-conditioned driving scene generation. In *NeurIPS*, 2024. 1

[86] Yunsong Zhou, Naisheng Ye, William Ljungbergh, Tianyu Li, Jiazhi Yang, Zetong Yang, Hongzi Zhu, Christoffer Petersson, and Hongyang Li. Decoupled diffusion sparks adaptive scene generation. In *ICCV*, 2025. 1, 6