Optimization-Guided Diffusion for Interactive Scene Generation

Shihao Li[1,2,3]✉, Naisheng Ye[2], Tianyu Li[2†], Kashyap Chitta[4,5], Tuo An[3], Peng Su[6], Boyang Wang[1], Haiou Liu[1], Chen Lv[3], and Hongyang Li[2]

[1] Beijing Institute of Technology, Beijing, China
[2] OpenDriveLab, The University of Hong Kong, Hong Kong, China
[3] Nanyang Technological University, Singapore
[4] KE:SAI – Kyutai ELLIS Scalable Autonomous Intelligence, Tübingen, Germany
[5] NVIDIA Research, Germany
[6] Yinwang Intelligent Tech. Co. Ltd., Shenzhen, China
lishihao.shawn@gmail.com
https://opendrivelab.com/OMEGA

Abstract. Realistic and diverse multi-agent driving scenes are crucial for evaluating autonomous vehicles, but safety-critical events which are essential for this task are rare and underrepresented in driving datasets. Data-driven scene generation offers a low-cost alternative by synthesizing complex traffic behaviors from existing driving logs. However, existing models often lack controllability or yield samples that violate physical or social constraints, limiting their usability. We present OMEGA, an optimization-guided, training-free framework that enforces structural consistency and interaction awareness during diffusion-based sampling from a scene generation model. OMEGA re-anchors each reverse diffusion step via constrained optimization, steering the generation towards physically plausible and behaviorally coherent trajectories. Building on this framework, we formulate ego–attacker interactions as a game-theoretic optimization in the distribution space, approximating Nash equilibria to generate realistic, safety-critical adversarial scenarios. Experiments on nuPlan and Waymo show that OMEGA improves generation realism, consistency, and controllability, increasing the ratio of physically and behaviorally valid scenes from 32.35% to 72.27% for free exploration capabilities, and from 11% to 80% for controllability-focused generation. Our approach can also generate $5\times$ more near-collision frames with a time-to-collision under three seconds while maintaining the overall scene realism.

Keywords: Autonomous Driving · Interactive Scene Generation · Diffusion Models · Constrained Optimization · Guided Sampling

1 Introduction

Generating realistic multi-agent driving scenes is essential for the development and evaluation of autonomous vehicles. However, existing rule-based or physics-based simulators [11, 30] fall short: they can reproduce simple traffic patterns but often yield

†: Project Lead.

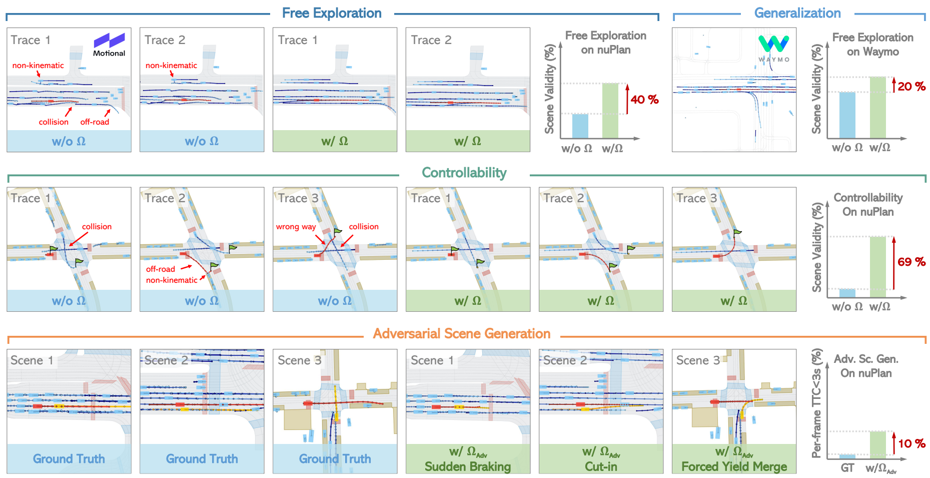

Fig. 1: OMEGA (Ω)**.** Our training-free, optimization-guided sampler can be dropped into existing diffusion-based driving scene generators to provide *significantly more realistic, controllable, and interactive* traffic scenarios. **Top.** Our guidance produces diverse and realistic future scenarios from the same historical initialization, increasing the rate of physically and behaviorally valid scenes on both nuPlan (used during the base model's training) and Waymo (zero-shot deployment) compared to unguided baselines. **Middle.** Controllability-conditioned generation synthesizes scenes where vehicles must reach predefined goal points (green flags) for one or multiple target agents, achieving precise behavioral control while maintaining natural interactions. **Bottom.** Finally, we enable adversarial scene generation, adaptively producing diverse attack scenarios without explicitly specifying attacker trajectories. Color gradients indicate temporal progression along each trajectory: **ego** (red gradient), **attacker** (yellow gradient), and **other vehicles** (blue gradient).

unrealistic or overly scripted interactions. Real-world datasets [1, 9] also offer limited coverage, particularly for safety-critical long-tail events that are crucial for assessing robustness yet occur too infrequently to capture at scale [29]. These shortcomings highlight the need for a high-fidelity driving scene generator capable of producing realistic, diverse, and controllable multi-agent interactions, especially those involving rare behaviors, which has become an indispensable requirement for efficient and safe evaluation of autonomous driving systems [3, 26, 27, 31, 37].

To meet these requirements, researchers have turned to data-driven scene generation, which enables models to learn complex agent behaviors directly from large-scale driving datasets [13, 19, 32]. Among these, diffusion models have shown remarkable potential in generating realistic multi-agent trajectories and coherent scene layouts, owing to their strong capacity to model high-dimensional, multi-modal distributions and produce samples consistent with empirical data [22, 25]. Despite this success, existing diffusion-based generators face two key limitations.

First, while the learned denoising network implicitly encodes physical and social constraints from data during training, such structure often degrades during inference. As noise is progressively removed, small inconsistencies can accumulate across steps, leading to trajectories that violate kinematic feasibility, social compliance, or interaction logic. These physically or behaviorally inconsistent samples reduce the credibility of generated scenes and limit their utility for downstream safe evaluation.

Second, long-tail and adversarial interactions lie in low-density regions of naturalistic datasets and are therefore underrepresented during training. Consequently, learned models gravitate toward dominant motion patterns and struggle to directly sample rare but safety-critical interactions. Prior diffusion methods introduce guidance via reinforcement-learned classifiers [39], manually designed behavior-shaping objectives [40, 44], or pre-specified goal conditions [24, 45]. These strategies either require retraining auxiliary networks or rely on case-by-case guidance functions that demand domain expertise, which limits scalability and generalization across scenarios.

To address these limitations, we introduce OMEGA, a training-free framework that incorporates structural constraints and strategic intent into diffusion-based driving scene generation. OMEGA formulates each reverse diffusion step as a conditional sampling process anchored on the estimated clean sample and performs numerical optimization within a KL-bounded trust region to re-anchor the mean of the reverse transition distribution. This optimization-guided formulation progressively steers the reverse Markov chain toward physically consistent and behaviorally coherent trajectories, improving structural fidelity without retraining or altering the backbone diffusion model.

To improve interaction realism, we introduce a phase-aligned guidance schedule that progressively transitions from feasible motion formation (per agent) to reactive coordination (among agents) across two denoising phases. Full-horizon coarse denoising under agent-wise structural priors first establishes a globally plausible motion layout, while fine-grained, time-indexed denoising resolves inter-agent interactions to enhance local reactivity while preserving physical feasibility.

Building on this foundation, we further introduce $\text{OMEGA}_{\text{Adv}}$, a sensitivity-enhanced adversarial generator inspired by [15, 35]. It models attacker–ego interactions as a distributional optimization game. The ego agent seeks to maintain natural, feasible behavior, while the attacker, using a sensitivity term that predicts how the ego will react to impending collision-avoidance constraints, seeks to maximally perturb the ego within the learned data distribution. This formulation enables the adaptive generation of realistic yet challenging driving scenarios without retraining any components or manually prescribing attacker trajectories or target points. Our contributions include:

- **Optimization-guided diffusion sampling.** A training-free method that re-anchors each reverse step via constrained optimization, enforcing structural consistency and stable diffusion guidance.
- **Phase-aligned guidance scheduling.** A stage-wise denoising and guidance scheme that shifts from macro-level plausibility to fine-grained reactivity, improving interaction fidelity and coordinated multi-agent scene evolution.
- **Sensitivity-enhanced adversarial generation.** A game-thoretic formulation for ego–attacker interactions yielding realistic yet challenging scenarios.
- **Comprehensive empirical results.** OMEGA reduces collisions and off-road rates while improving kinematic feasibility over baselines. It raises scene validity on nuPlan by about 40 points and on Waymo (zero-shot) by about 20 points, and delivers a further 69-point gain in controllability under user-specified intents, mitigating prior validity limitations and bringing diffusion-based scene generation closer to practical

integration. It is also capable of generating $5\times$ more near-collision frames under the adversarial scene generation setting.

2 Related Work

Diffusion models for conditioned generation. A spectrum of generative architectures has been explored to address driving scenario generation, encompassing autoregressive approaches [19, 32] and diffusion-based frameworks [7, 25, 36, 45]. The latter offers greater flexibility for joint distribution modeling and supports conditional synthesis through guidance mechanisms. Several guidance strategies have been proposed for trajectory- and scene-level diffusion models. Some approaches, akin to diffusion forcing [4], interpret Gaussian corruption as continuous partial masking, allowing low-noise tokens to guide reconstruction for macro-level behavior control [23, 45] or temporally structured rollouts [25]. While effective for coarse behavioral shaping, such noise modulation provides limited control over fine-grained motion regulation and interaction consistency. Other approaches introduce guidance during denoising through score or state corrections. CTG [43, 44] approximates conditional scores using predicted reverse-step means to enable flexible conditional control, while DPS-style methods [8, 21, 24, 42] further align reverse diffusion updates with clean-space objectives by propagating gradients through the denoiser. However, this guidance relies on local linearization of the nonlinear noise–state mapping and may produce inaccurate update directions. Geometry-constrained approaches such as GHC [25] instead enforce feasibility through projection or clipping, introducing discontinuities that may compromise the smoothness and distributional consistency of the generated trajectories. In contrast, OMEGA integrates noise scheduling and reverse-step guidance in a phase-specific framework, decoupling structural layout formation from fine-grained reactive refinement. By re-anchoring each reverse transition distribution via KL-bounded constrained optimization in the clean space, OMEGA provides more reliable structural control and interaction consistency.

Safety-critical scenario generation. Generating high-fidelity, safety-critical driving scenarios is paramount for the robust training and evaluation of autonomous driving systems [5, 20, 28]. Certain approaches employ adversarial learning [14, 41], which trains surrounding agents to adversarially interact with the ego vehicle. Gradient-based perturbation methods, including KING [16] and AdvSim [38], modify background trajectories, or re-simulate sensor data, to induce critical states while preserving physical plausibility. AdvDiffuser [39] injects gradients into the denoising process from an auxiliary collision reward model, steering the diffusion updates toward adversarial trajectories. Nexus [45] leverages corner-case finetuning and manually specified goal-conditioned inpainting to steer selected agents toward predefined outcomes. In contrast to these approaches, our method, $\text{OMEGA}_{\mathrm{Adv}}$, generates scenarios by formulating a distributional game in the optimization guided denoising process, enabling realistic yet safety-critic scenario generation.

3 Method

3.1 Preliminaries

Diffusion models. Diffusion models [10, 17, 33, 34] are likelihood-based generative models that map a simple Gaussian prior to complex data distributions through sequential denoising. Given $x_0 \sim q(x_0)$, the *forward diffusion* adds Gaussian noise with variance $\beta_t \in (0, 1)$ at each step to obtain a sequence of intermediate states $x_1, \ldots, x_T$. The analytic marginal distribution of conditioned on x_0 is:

$$
\begin{aligned}
q(x_t \mid x_0) &= \mathcal{N}\big(\sqrt{\bar{\alpha}_t}\, x_0,\ (1 - \bar{\alpha}_t) I\big), \\
x_t &= \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1 - \bar{\alpha}_t}\, \epsilon.
\end{aligned}
\tag{1}
$$

where $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$, and $\epsilon \sim \mathcal{N}(0, I)$.

Starting from pure noise $x_T \sim \mathcal{N}(0, I)$, the *reverse process* reconstructs data from noise and is modeled as a Gaussian transition:

$$
p_\theta(x_{t-1} \mid x_t) = \mathcal{N}\big(\mu_\theta(x_t, t),\ \Sigma_\theta(x_t, t)\big), \tag{2}
$$

where in DDPMs [17] the variance Σ_θ is fixed as $\sigma_t^2 I$, and the network learns to predict the noise term $\epsilon_\theta(x_t, t)$ that perturbed x_0 to obtain x_t. After training, samples are generated by iteratively applying this reverse transition:

$$
x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}}\, \epsilon_\theta(x_t, t) \right) + \sigma_t z, \tag{3}
$$

where $z \sim \mathcal{N}(0, I)$ is standard Gaussian noise and $t = T, \ldots, 1$. Here, β_t denotes the noise schedule, $\alpha_t = 1 - \beta_t$ is the retained signal coefficient, and $\bar{\alpha}_t$ is its cumulative product controlling the noise level.

Problem formulation. We represent a traffic scene as a spatiotemporal tensor $x \in \mathbb{R}^{A \times \mathcal{T} \times D}$ where A denotes the number of interacting agents, $\mathcal{T}$ the total number of physical timesteps, and D the state dimension including position, heading, velocity, and physical size. A binary validity mask $m \in \mathbb{B}^{A \times \mathcal{T}}$ indicates each agent's existence over time, accounting for dynamic entries and exits. Following prior works [25, 45], we formulate driving scene generation as a multi-agent inpainting problem over x. Given an inpainting mask $\tilde{m} \in \mathbb{B}^{A \times \mathcal{T} \times D}$ and its corresponding context values $\bar{x} = \tilde{m} \odot x$, the diffusion model learns to reconstruct the unobserved elements of x conditioned on both local spatiotemporal and global scene information. The given context may include arbitrary known elements of x, such as all agents' past trajectories or designated future goals for specific agents. The effective target region for generation corresponds to valid but unobserved elements $m \odot (1 - \tilde{m})$, representing future state of existing agents or newly introduced ones. Within this formulation, the diffusion model iteratively refines the scene tensor through denoising, progressively completing the partially observed scene toward a coherent multi-agent configuration.

3.2 Optimization-Guided Diffusion Sampling

Anchored reverse transition. In the standard diffusion sampling process, the model predicts the injected noise $\epsilon_\theta(x_t, t)$ at each step, from which an approximate clean estimate corresponding to the noisy state x_t can be obtained via Eq. (1):

$$\tilde{x}_0 = \frac{x_t - \sqrt{1-\bar{\alpha}_t}\,\epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}}. \tag{4}$$

Substituting $\tilde{x}_0$ into the reverse kernel, the posterior distribution of x_{t-1} given x_t can be expressed as:

$$p_\theta(x_{t-1} \mid x_t) = \mathcal{N}\big(x_{t-1};\, A_t\tilde{x}_0 + C_t x_t,\, \sigma_t^2 I\big), \tag{5}$$

where $A_t = \frac{(1-\alpha_t)\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_t}$ and $C_t = \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}$ are deterministic coefficients determined by the variance schedule. Since x_t, A_t, and C_t are fixed at each step, the conditional distribution of x_{t-1} is fully characterized by the model-estimated clean sample $\tilde{x}_0$. For convenience, we denote this Gaussian transition as $Q_t(\tilde{x}_0)$, indicating that the reverse kernel is anchored on $\tilde{x}_0$ while its variance and linear coefficients are fixed by (x_t, t).

Accordingly, Eq. (3) can be rewritten as:

$$x_{t-1} = \underbrace{A_t\tilde{x}_0}_{\text{regression term}} + \underbrace{C_t x_t}_{\text{inertia term}} + \underbrace{\sigma_t z}_{\text{noise term}}. \tag{6}$$

This update can be decomposed into three interpretable components: (i) a **regression term** ($A_t\tilde{x}_0$) that pulls the sample toward the model-predicted clean state, (ii) an **inertia term** ($C_t x_t$) that preserves continuity with the current noisy state and stabilizes transitions across steps, and (iii) a **noise term** ($\sigma_t z$) that maintains sample diversity through controlled stochasticity. Together, these components describe a reverse process that iteratively refines x_t toward the predicted clean state $\tilde{x}_0$, ultimately yielding a sample consistent with the learned data distribution. This perspective highlights that each reverse diffusion step constitutes a *conditional sampling process anchored on the predicted clean sample*, with $\tilde{x}_0$ serving as the denoising target that guides the evolution of the Markov chain.

Optimization-guided re-anchoring. While the predicted clean sample $\tilde{x}_0$ defines the local denoising direction inferred from the learned dynamics, it is derived purely through data-driven regression and may therefore violate structural or behavioral priors, gradually drifting away from the manifold of physically and behaviorally plausible samples.

To explicitly regularize this evolution, OMEGA introduces an optimization-guided refinement of the anchor (Fig. 2), redefining each reverse transition as a new distribution $P_t(\hat{x}_0)$ whose mean is adaptively adjusted within a bounded divergence from the original reverse kernel:

$$\begin{aligned} &P_t(\hat{x}_0) = \mathcal{N}(A_t\hat{x}_0 + C_t x_t,\, \sigma_t^2 I), \\ &\text{s.t.} \quad \mathrm{KL}\big(P_t(\hat{x}_0) \,\|\, Q_t(\tilde{x}_0)\big) \le \kappa_t. \end{aligned} \tag{7}$$

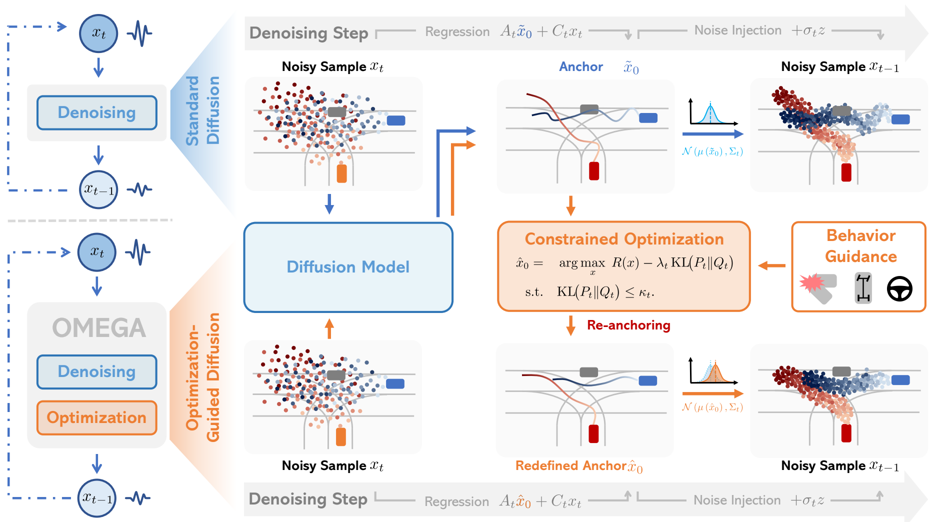

Fig. 2: Standard diffusion sampling (top) vs. our optimization-guided OMEGA (bottom). We re-anchor each reverse diffusion step via constrained optimization within a KL-bounded trust region, steering the Markov chain towards behaviorally coherent samples.

The optimized anchor $\hat{x}_0$ is obtained by solving a constrained maximization problem:

$$\begin{aligned} \hat{x}_0 = \arg\max_x \Big[\lambda_t R(x) - \mathrm{KL}\big(P_t(x) \,\|\, Q_t(\tilde{x}_0)\big)\Big], \\ \text{s.t.} \quad \mathrm{KL}\big(P_t(x) \,\|\, Q_t(\tilde{x}_0)\big) \le \kappa_t. \end{aligned} \tag{8}$$

Here, $R(x)$ serves as a differentiable objective that encodes structural and behavioral preferences guiding the generation process, while the KL term defines a trust region that regularizes deviation from the pretrained reverse transition. As illustrated by the toy example in Fig. 3, this constraint prevents excessive drift or mode collapse, stabilizing the denoising dynamics and preserving consistency with the learned distribution. Through this balance, the sampler implicitly shifts the mean of the reverse transition toward constraint-consistent regions, adapting its trajectory while maintaining distributional coherence throughout the reverse process.

OMEGA guides the denoising process through distributional re-anchoring, softly steering each reverse step toward constraint-consistent regions while limiting per-step drift and mitigating intermediate-step collapse. By iteratively optimizing the anchor $\hat{x}_0$ and resampling from $P\left(\hat{x}_0\right)$, the process progressively re-aligns the reverse Markov chain with both the model's learned data distribution and the imposed structural priors, enabling structural yet distributionally coherent generation.

Equivalent euclidean formulation. Since the anchor $\hat{x}_0$ directly corresponds to the joint future trajectories of all agents in the scene, it admits a physically meaningful interpretation in the optimization space. We represent $R(x)$ as a structured objective $r(x)$ equipped with equality and inequality constraints $h(x) = 0$ and $g(x) \le 0$, encoding

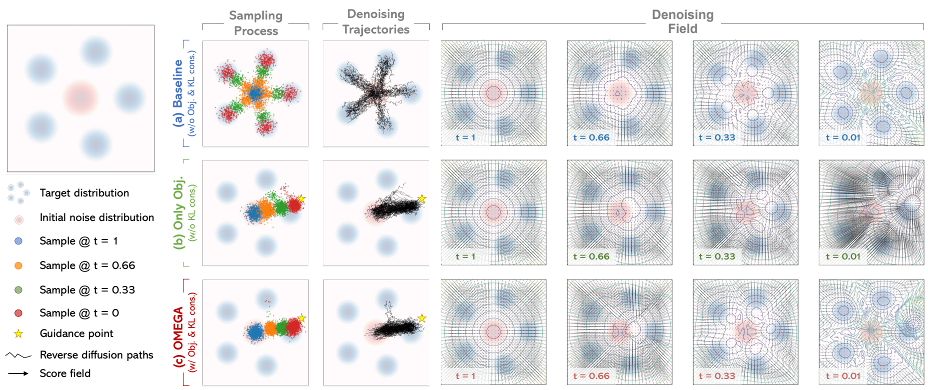

Fig. 3: Toy example illustrating the effect of OMEGA on generative sampling dynamics. We adopt a simple two-dimensional generation setup from [18] to visualize how our method influences the sampling process. (a) Without guidance, the diffusion model approximates the target data distribution and recovers its multimodal structure. (b) Using only the guidance objective $R(x)$ pulls samples excessively toward the guidance point, thereby distorting the local sampling geometry. (c) Combining the guidance objective with a KL constraint yields controlled adaptation toward the guidance point while suppressing excessive distributional drift. Details are provided in the supplementary material.

different levels of structural and behavioral priors. Here, $r(x)$ captures trajectory-level preferences such as smoothness, $h(x) = 0$ enforces equality relations such as kinematic feasibility, and $g(x) \leq 0$ defines inequality conditions such as collision avoidance or boundary adherence.

By substituting the Gaussian forms of $P_t(x)$ and $Q_t(\tilde{x}_0)$ into Eq. (8), the optimization can be explicitly written in the Euclidean domain as:

$$\begin{aligned} \hat{x}_0 = \arg\max_{x} \;& \lambda_t\, r(x) - \frac{A_t}{2\sigma_t^2}\, \|x - \tilde{x}_0\|_2^2, \\ \text{s.t.} \;& \|x - \tilde{x}_0\|_2 \leq \sqrt{2\kappa_t}\, \frac{\sigma_t}{|A_t|}, \\ & h(x) = 0, \quad g(x) \leq 0. \end{aligned} \tag{9}$$

This equivalent formulation shows that the KL-divergence constraint in Eq. (8) translates into an adaptive Euclidean trust region, whose radius scales proportionally with the diffusion variance σ_t. The trust region bounds the magnitude of each refinement step, while the guidance signal shapes its direction according to the local scene context. Consequently, early reverse steps with higher stochastic uncertainty allow the optimization to explore a broader set of feasible configurations, whereas subsequent steps at lower variance confine updates near the learned data manifold, thereby promoting stable convergence and preserving consistency with the learned generative dynamics.

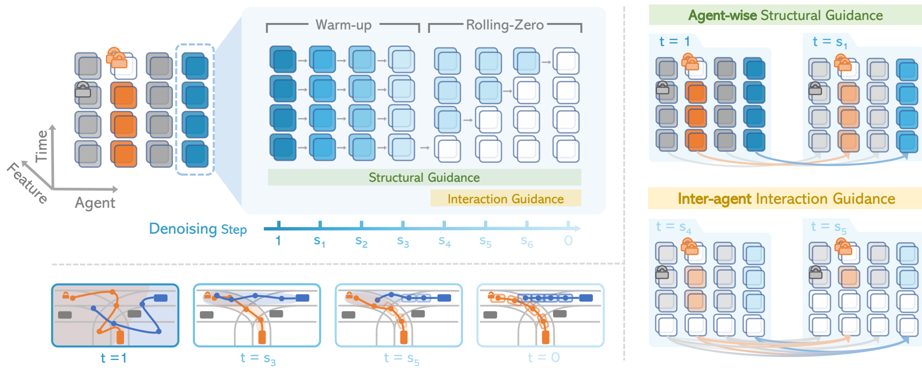

Fig. 4: Illustration of phase-aligned guidance schedule. Darker blue means higher noise, lighter blue means lower noise. The green and yellow bars denote phase-specific guidance activation: agent-wise structural guidance is applied throughout denoising to refine each agent under its own motion and scene constraints, while inter-agent interaction guidance is additionally activated during Rolling-Zero to condition updates on the latest denoised states of other agents.

3.3 Phase-Aligned Guidance Schedule

Realistic multi-agent scene generation requires jointly achieving globally plausible motion evolution and fine-grained local reactivity. This entails satisfying heterogeneous objectives for agent-wise motion feasibility and inter-agent coordination. Uniformly enforcing these objectives throughout denoising essentially casts the reverse process as solving a high-dimensional, tightly coupled problem over multiple agents and time steps. More critically, before basic feasible motion structure is established, interaction objectives can be ill-conditioned due to unreliable relative configurations across agents. Enforcing them jointly with objectives for motion feasibility in this regime can therefore yield conflicting guidance signals and unnecessary optimization overhead.

To address this issue, we propose a phase-aligned guidance schedule, which integrates a two-phase denoising schedule with phase-specific guidance, as shown in Fig. 4. The denoising schedule, consisting of *Warm-up* and *Rolling-Zero*, decomposes scene generation into macro-level layout formation and fine-grained, time-indexed interaction refinement. The guidance objectives are then activated in a phase-aligned manner to reinforce their distinct roles in feasible motion formation and reactive coordination.

Warm-up phase. The Warm-up phase starts from the fully corrupted scene state and performs full-sequence denoising over the entire future horizon under a gradually decreasing noise schedule. This global optimization establishes macro-level spatial organization and dynamic plausibility, ensuring smooth and kinematically coherent motion throughout the horizon. By treating the temporal sequence as a unified denoising target, Warm-up suppresses autoregressive drift and stabilizes long-horizon trends. It then halts in a low-noise regime, retaining controlled residual uncertainty and providing a well-posed initialization for subsequent refinement.

Rolling-Zero phase. Operating in the low-noise regime produced by Warm-up, the Rolling-Zero phase progressively removes the residual noise through a time-indexed denoising schedule. At each step, one frame is refined conditioned on the progressively current denoised history and partially masked future states. This temporally localized process enables fine-grained adaptation to evolving inter-agent interactions while preserving the globally coherent structure established in Warmup, thereby gradually restoring frame-level responsiveness.

Phase-specific guidance. Guidance is scheduled across the two phases. During Warm-up, it is restricted to agent-wise structural objectives, including motion feasibility, road compliance, state consistency, and smoothness, to form a globally plausible motion layout and coarse behavior intent. During Rolling-Zero, inter-agent interaction guidance is activated in a time-indexed manner and evaluated on the latest denoised scene state, enabling timely responses to evolving multi-agent motions. This phase-aligned design reduces joint optimization burden and naturally supports parallel computation.

3.4 Sensitivity-Enhanced Adversarial Generation

Building upon the OMEGA formulation, $\text{OMEGA}_{\text{Adv}}$ extends it to adversarial scene generation by formulating a sensitivity-enhanced distributional game between the ego vehicle e and an attacker a. While the ego aims to preserve feasibility and realism within the learned generative manifold, the attacker seeks to induce maximal perturbations that remain distributionally consistent, thereby exposing safety-critical yet physically plausible interactions. This formulation enables adversarial scenario synthesis directly within the learned generative space, without retraining or external trajectory scripting.

Following this distributional game setup, both participants perform constrained optimization: the ego strives to maintain realistic feasibility:

$$\begin{aligned} \hat{x}_0^e = &\arg\max_{x^e} \; J^e(x^e) \\ \text{s.t.} \quad &\|x^e - \tilde{x}_0^e\|_2 \le \rho_t, \; h(x^e) = 0, \; g(x^e) \le 0, \\ &\gamma_e(x^e, x^a) \le 0, \end{aligned} \tag{10}$$

while the attacker perturbs ego behavior by solving:

$$\begin{aligned} \hat{x}_0^a = &\arg\max_{x^a} \; J^a(x^a) - \alpha J^e(x^e) \\ \text{s.t.} \quad &\|x^a - \tilde{x}_0^a\|_2 \le \rho_t, \; h(x^a) = 0, \; g(x^a) \le 0, \\ &\gamma_a(x^a, x^e) \le 0. \end{aligned} \tag{11}$$

Here, J^e and J^a share the same structural form as Eq. (9), promoting consistency with the learned generative manifold, while $\alpha > 0$ balances the attacker's aggressiveness against realism. The pairwise safety constraint $\gamma(x^e, x^a)$ couples the two subproblems, inducing a noncooperative differential game where each player's feasible set depends on the other's strategy and is activated only within their respective responsibility regions.

Sensitivity-enhanced iterative best response. Directly solving the joint optimization in Eqs. (10)-(11), which constitutes a noncooperative differential game, is computationally prohibitive. To approximate its Nash equilibrium efficiently, we adopt a sensitivity-enhanced iterative best response (SE-IBR) scheme [35]. Starting from the model-predicted anchors $\tilde{x}_0^e$ and $\tilde{x}_0^a$ as initial strategies, the ego and attacker alternately update their clean anchors as best responses to each other, iterating until convergence.

Let $x^{a(l-1)}$ denote the attacker's previous anchor and $x^{e(l)}$ the ego's current optimized response. Since the ego's optimal value $J^{e*}(x^a)$ implicitly depends on the attacker's decision x^a, SE-IBR characterizes its local sensitivity around $x^{a(l-1)}$ via a first-order Taylor approximation:

$$J^{e*}(x^a) \approx J^{e*}\big(x^{a(l-1)}\big) + \left.\frac{\mathrm{d}J^{e*}}{\mathrm{d}x^a}\right|_{x^{a(l-1)}} (x^a - x^{a(l-1)}). \tag{12}$$

Using the Karush-Kuhn-Tucker (KKT) optimality conditions of the ego's constrained problem (10), the sensitivity of J^{e*} with respect to the attacker's decision can be approximated around the current iterate as:

$$\left.\frac{\mathrm{d}J^{e*}}{\mathrm{d}x^a}\right|_{x^{a(l-1)}} \approx -\mu^{e(l)} \left.\frac{\partial \gamma_e}{\partial x^a}\right|_{(x^{a(l-1)}, x^{e(l)})}, \tag{13}$$

where $\mu^{e(l)} \geq 0$ is the KKT multiplier for ego's active collision constraint $\gamma_e(x^a, x^e) \leq 0$. Substituting Eqs. (12)-(13) yields the attacker's sensitivity-enhanced update:

$$\begin{aligned} \hat{x}_0^a \approx \ & \arg\max_{x^a} \ J^a(x^a) + \alpha\mu^{e(l)} \left.\frac{\partial \gamma_e}{\partial x^a}\right|_{(x^{a(l-1)}, x^{e(l)})} x^a \\ & \text{s.t.} \quad \text{(same constraints as (11)).} \end{aligned} \tag{14}$$

When the ego's avoidance constraint becomes active (i.e. $\mu^{e(l)} > 0$), the attacker acquires a directional incentive proportional to $\partial\gamma_e/\partial x^a$, which encourages attacker motions that decrease the signed-distance in the ego–attacker interaction space. Intuitively, this steers the attacker toward maneuvers that effectively reduce the ego's feasible maneuvering set and thus elicit a reactive avoidance response from the ego, while the KL-based trust-region preserves distributional plausibility under the pretrained model. Because the collision-avoidance constraint only activates within the ego's responsibility region, this design indirectly ensures the generated adversarial behaviors correspond to plausible and meaningful interactions rather than irrelevant collisions (e.g., passive rear-endings). To raise the probability of entering this active regime in practice, we bias attacker route initialization during the Warmup phase toward spatial corridors that intersect the ego's responsibility region. Once the collision-avoidance constraint is activated, the first-order sensitivity term governs the attacker's subsequent refinements, yielding targeted yet distributionally consistent adversarial behaviors. Detailed sensitivity derivations are provided in the Appendix.

4 Experiments

4.1 Setup and Protocol

We employ the Nexus model [45], a diffusion-based driving scene generator pretrained on the nuPlan [2] dataset, and apply our inference procedure directly on its outputs without any additional training or fine-tuning, named Nexus-Ω. We compare our method with

Table 1: Free-exploration scene generation results on nuPlan and zero-shot evaluation on Waymo. Oracle refers to original logs in the dataset. N-Dist. (nearest-agent distance distribution JSD, 10^{-3}), L-Dev. (lateral deviation distribution JSD, 10^{-3}), A-Dev. (angular deviation distribution JSD, 10^{-3}), Spd. (speed distribution JSD, 10^{-3}), P-Ag./P-Sc. (per-agent/per-scene rate), and overall Scene Valid Rate (no collision/off-road/kinematic violation). ↓ indicates lower is better; ↑ indicates higher is better.

Dataset	Method	**Distributional JSD**				**Col.(%)**		**Off-road(%)**		**K.Feas.(%)**		**Valid(%)**
		N-Dist. ↓	L-Dev. ↓	A-Dev. ↓	Spd. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↑	P-Sc. ↑	P-Sc. ↑
nuPlan	Oracle	0.000	0.000	0.000	0.000	0.38	13.62	0.00	0.00	97.40	81.22	71.67
	D. Policy [6]	0.701	0.056	0.423	0.369	1.19	36.58	4.37	39.47	98.46	87.18	34.49
	SceneD. [25]	0.933	0.046	0.395	**0.234**	1.19	37.80	4.84	41.68	98.48	86.96	33.27
	Nexus [45]	1.162	0.041	0.461	1.018	1.41	43.12	4.70	40.13	98.72	88.63	32.35
	Nexus-Ω	**0.203**	**0.025**	**0.081**	0.333	**0.19**	**17.36**	**0.91**	**10.39**	**99.19**	**91.91**	**72.27**
Waymo	Oracle	0.000	0.000	0.000	0.000	0.22	7.47	0.00	0.00	99.77	96.33	89.33
	D. Policy [6]	0.034	0.026	0.081	0.229	0.35	19.23	1.04	11.35	98.86	86.74	64.55
	SceneD. [25]	0.077	0.022	0.148	0.249	0.39	22.94	1.22	12.18	98.81	86.57	61.01
	Nexus [45]	0.081	0.022	0.154	0.251	0.38	22.65	1.23	12.30	98.76	86.08	61.71
	Nexus-Ω	**0.019**	**0.015**	**0.050**	**0.198**	**0.18**	**11.43**	**0.37**	**4.32**	**99.40**	**92.65**	**81.95**

recently available and faithfully reproduced diffusion-based scene generation approaches trained on the nuPlan dataset under multiple settings. Details of our experimental settings and baselines are provided in the Appendix.

4.2 Comparison to State of the Art

In this section, we assess the method's realism and physical plausibility, generalizability to unseen environments, and goal-conditioned controllability.

Free exploration evaluation. The model generates multiple plausible futures conditioned on the history states in this setting. As shown in Table 1, our approach, denoted as Nexus-Ω, delivers consistent gains in both realism and physical plausibility over prior diffusion-based generators. On nuPlan, it substantially reduces failure cases, cutting scene-level collisions by **25.76** points and off-road events by **29.74** points, while improving kinematic feasibility by **3.3** points. These gains result in a **72.27%** valid rate, nearly doubling Nexus' overall performance. When evaluated zero-shot on the unseen Waymo [12] dataset, Nexus-Ω maintains its advantage, achieving a valid rate of **81.95%**, **+20.24** points higher than Nexus, demonstrating strong generalizability. To further assess the generality of OMEGA and its effect on distributional realism, we apply it to both Nexus and VBD [21] and evaluate them on the Waymo Open Sim Agents *val* set, with Nexus evaluated zero-shot. As shown in Table 2, OMEGA improves both generators. The larger gain for Nexus highlights the greater benefits of guidance for interaction and map-based realism under zero-shot shift.

Goal-conditioned controllability. In this setting, we specify an explicit goal point for a target agent at its final timestep and adjust the inpainting mask $\tilde{m}$ accordingly. The model is expected to generate a feasible scene completion that retains overall realism under the specified inpainting condition. Unlike prior approaches that often enforce goal reaching

Table 2: Distributional realism under the Waymo Open Sim Agents evaluation protocol. We compare Nexus and VBD with their respective Ω-enhanced variants.

Method	Kinematics ↑				Interaction ↑			Map-Based ↑		Realism
	Lin. Spd.	Lin. Acc.	Ang. Spd.	Ang. Acc.	Obj. Dist.	Col.	TTC	Road Dist.	Offroad	Meta ↑
Nexus	0.253	0.377	0.324	0.611	0.290	0.763	0.833	0.573	0.683	0.609
Nexus-Ω	0.246	0.321	0.329	0.598	0.302	0.884	0.835	0.609	0.919	0.700
VBD	0.353	0.332	0.437	0.563	0.341	0.943	0.820	0.631	0.895	0.723
VBD-Ω	0.352	0.353	0.441	0.567	0.358	0.951	0.823	0.640	0.898	0.730

Table 3: Comparison on goal-conditioned controllability. Each method is evaluated by setting multiple user-defined goal points for the target vehicle to test controllability.

Method	Distributional JSD				Col.(%)		Off-road(%)		K.Feas.(%)		Valid(%)
	N-Dist. ↓	L-Dev. ↓	A-Dev. ↓	Spd. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↑	P-Sc. ↑	P-Sc. ↑
D. Policy [6]	0.382	0.706	0.047	73.245	0.81	40.00	3.71	46.00	96.32	58.00	20.00
SceneD. [25]	0.526	1.405	0.033	75.818	0.88	46.00	4.21	40.00	96.63	59.00	17.00
Nexus [45]	0.594	1.618	0.037	76.187	1.20	51.00	5.22	51.00	96.44	55.00	11.00
Nexus-Ω	**0.222**	**0.244**	**0.012**	**5.151**	**0.10**	**12.00**	**0.44**	**7.00**	**99.37**	**91.00**	**80.00**

through abrupt trajectory deformation, leading to large lane deviation, off-road or dynamically infeasible motions, Nexus-Ω integrates the goal constraint naturally during the denoising process, allowing smooth adherence to the target without compromising feasibility. As shown in Table 3, OMEGA reduces the scene-level collision rate (**-39.00**) and the off-road rate (**-44.00**), and boosts kinematic feasibility (**+36.00**) over Nexus. This improvement results in an increase of **+69.00** points, nearly an order-of-magnitude gain over the baselines. Overall, OMEGA enables precise goal controllability while maintaining natural multi-agent coordination and physical plausibility.

4.3 Adversarial Scenario Generation

Generic adversarial generation. To characterize the strength and realism of the generated adversarial long-tail interactions, we compare Nexus-Ω_{Adv} with multiple baselines. As shown in the left and right panels of Fig. 5, our guidance strategies produce substantially more interactive scenes than the GT distribution and other Nexus baselines. This is evidenced by consistently lower mean TTC and higher proportions of ego TTC below 1–3 seconds, indicating increased risk exposure for the ego vehicle. Correspondingly, ego trajectory statistics, including mean acceleration and jerk increase, suggest that generated scenarios trigger stronger evasive or aggressive responses. Importantly, despite the increased attack intensity, off-road rates and kinematic feasibility remain acceptable (lower off-road, higher kinematic feasibility), and the Ego NC metric confirms that the generated collisions are more meaningful–reflecting realistic, ego-responsible interactions that provide greater training and evaluation value. Together, these results demonstrate that our method substantially increases scenario aggressiveness while preserving physical plausibility.

Fig. 5: Adversarial scene generation results. (left) Comparison between oracle (GT), Nexus variants, including: Nexus finetuned on adversarial scenarios (Nexus-FT); Nexus with goal-attacking condition (Nexus-GC); Nexus-CTG with goal-attacking cost (Nexus-CTG$_{\mathrm{Adv}}$). Ego Risk: mean time to collision per frame (TTC, s), percentage of TTC $< 1, 2, 3\,s$. Intensity of Ego Motion: mean accleration (Acc, m^2/s), jerk (m^3/s). Ego Non-responsible Collision (Ego NC, %). **(right)** Nexus-Ω_{Adv} produces more samples with short TTC and low speed, indicating the generation of riskier scenarios.

Method	Ego Risk				Ego Motion		Ego NC	Off-road	K. Feas.
	TTC ↓	<1s ↑	<2s ↑	<3s ↑	Acc. ↑	Jrk. ↑	P-Sc. ↓	P-Sc. ↓	P-Sc. ↑
Oracle	4.904	0.328	0.949	2.212	0.360	0.307	0.97	0.00	81.22
Nexus-FT	4.695	3.239	5.453	7.214	0.355	0.477	9.91	39.77	87.08
Nexus-GC	4.631	3.462	6.179	8.677	0.631	1.384	10.98	45.14	45.16
Nexus-CTG$_{\mathrm{Adv}}$	4.575	3.541	6.254	8.747	0.794	**2.145**	7.12	26.77	85.64
Nexus-Ω_{Adv}	**4.516**	**3.450**	**7.721**	**11.599**	**0.859**	1.972	**5.42**	**15.82**	**88.98**

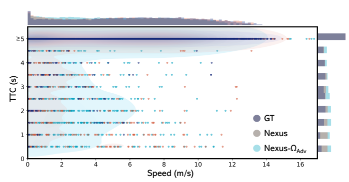

Table 5: Comparison of different guidance strategies and **ablation study**. Results under the *Free Exploration* setting.

Method	Distributional JSD				Col.(%)		Off-road(%)		K.Feas.(%)		Valid(%)
	N-Dist. ↓	L-Dev. ↓	A-Dev. ↓	Spd. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↓	P-Sc. ↓	P-Ag. ↑	P-Sc. ↑	P-Sc. ↑
Nexus [45]	1.162	0.041	0.461	1.018	1.41	43.12	4.70	40.13	98.72	88.63	32.35
Nexus-GHC [25]	2.368	0.041	0.694	3.222	2.23	49.05	1.58	11.57	92.50	50.29	26.06
Nexus-CTG [44]	0.392	0.031	0.225	0.713	0.44	26.02	0.92	9.97	97.13	81.72	50.41
Nexus-DPS [42]	0.421	0.028	0.214	0.677	0.51	27.25	0.95	11.34	97.56	83.55	47.19
w/o Rolling-Zero	0.251	0.028	0.108	2.168	1.19	34.35	**0.57**	**6.64**	**99.69**	**96.72**	60.49
w/o Warm-up	23.180	0.178	113.874	184.708	8.18	76.90	45.01	98.95	3.15	0.19	0.00
w/o Guid. Obj.	2.241	0.094	0.248	1.285	1.37	41.38	4.66	39.72	98.49	85.69	33.47
w/o Phase Align.	0.705	0.046	1.132	3.061	0.68	31.54	1.20	13.03	97.07	75.73	52.05
w/o KL cons.	2.544	0.060	6.773	53.086	3.15	63.40	5.43	46.38	89.97	51.95	18.21
Nexus-Ω	**0.203**	**0.025**	**0.081**	**0.333**	**0.19**	**17.36**	0.91	10.39	99.19	91.91	**72.27**

Planner-specific adversarial generation. By injecting the ego planned trajectory into the game as a weak-noise condition, our framework enables planner-specific adversarial generation. As shown in Table 4, Nexus-Ω_{Adv} consistently increases risk exposure across different planners by shifting scene minimum TTC toward shorter values and raising ego-at-fault collision rates. Meanwhile, the non-at-fault collision rate remains low, suggesting planner-relevant breakdowns rather than uninformative passive collisions.

Table 4: Planner-in-the-loop adversarial generation results. Three ego planners, log replay (Replay), constant velocity (CV), and IDM, are evaluated on Log and Nexus-Ω_{Adv}. Scene-level minimum TTC and collision rates are reported, with collisions split into ego-at-fault (A.F.) and non-at-fault (N.A.F.).

Planner	Scenarios	Scene Min TTC (s)				Scene Collision (%)	
		<1s↑	<2s↑	<3s↑	Avg↓	A.F.↑	N.A.F.↓
Replay	Log	0.07	1.85	4.72	4.97	0.02	**0.97**
	Nexus-Ω_{Adv}	**21.31**	**25.52**	**29.14**	**4.64**	**19.50**	4.47
CV	Log	14.82	16.88	19.06	4.79	8.00	10.35
	Nexus-Ω_{Adv}	**23.17**	**26.98**	**29.87**	**4.61**	**17.25**	**6.34**
IDM	Log	6.51	8.83	14.04	4.89	1.34	8.98
	Nexus-Ω_{Adv}	**18.55**	**22.34**	**25.66**	**4.64**	**15.62**	**5.37**

4.4 Further Analysis

Comparison with other guidance methods. As shown in Tab. 5, Nexus-Ω achieves the best overall performance among the projection-style guidance method GHC [25], the conditional score approximation method CTG [44], and the prior clean-space guidance method DPS [42]. It substantially improves realism, while reducing collision and off-road rates and improving kinematic feasibility. These consistent gains translate into a markedly higher overall valid rate compared with prior guidance approaches. The results suggest that optimization-guided diffusion more effectively balances realism with constraint satisfaction, yielding physically consistent scenes.

Ablation study. As shown in Tab. 5, removing the Rolling-Zero phase leads to smoother but less responsive trajectories: the model performs well on-road but exhibits higher collision rates. Using only the Rolling-Zero phase (without Warm-up) results in severe cumulative autoregressive errors and degraded performance across all metrics. Retaining the two-phase noise schedule without guidance objectives yields only marginal gains over the base model, indicating that noise scheduling alone is insufficient. Applying guidance without phase-specific alignment retains only part of the scene-validity gain while requiring $7.85\times$ the inference time of the full OMEGA configuration, increasing from $4.96{\pm}3.92, \mathrm{s}$ to $38.94{\pm}39.79, \mathrm{s}$, highlighting the benefits of phase-aligned guidance for both generation quality and computational efficiency. Omitting the KL divergence constraint drastically reduces distributional realism and overall stability. The full OMEGA configuration achieves the best balance among realism, smoothness, and safety, yielding the highest valid rate.

5 Conclusion

We present OMEGA, a training-free, optimization-guided framework to enhance the fidelity, interactivity, and controllability of diffusion-based driving scene generation. A two-stage noise schedule further enhances interaction realism by combining global trajectory refinement with local frame-level responsiveness. OMEGA$_{\mathrm{Adv}}$ then produces realistic and challenging long-tail scenarios. A limitation of our current implementation is its runtime: optimizing an 8-second scene can take around 5 seconds. Parallelizing our per-agent optimization can reduce this latency, enabling the integration of OMEGA into closed-loop simulation frameworks. We will leave it as future work.

Acknowledgments

This work is supported by the China Scholarship Council and the National Natural Science Foundation of China (Grant Nos. 52172378 and 52302489). This work is in part supported by the JC STEM Lab of Autonomous Intelligent Systems funded by The Hong Kong Jockey Club Charities Trust.

We sincerely thank Yuxuan Hu, Yuhang Lu, and Haochen Liu for their valuable assistance, and we extend our appreciation to the members of OpenDriveLab for their continuous support throughout this study.

References

1. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuScenes: A multimodal dataset for autonomous driving. In: CVPR (2020)
2. Caesar, H., Kabzan, J., Tan, K.S., Fong, W.K., Wolff, E., Lang, A., Fletcher, L., Beijbom, O., Omari, S.: nuPlan: A closed-loop ml-based planning benchmark for autonomous vehicles. arXiv preprint arXiv:2106.11810 (2021)
3. Cao, W., Hallgarten, M., Li, T., Dauner, D., Gu, X., Wang, C., Miron, Y., Aiello, M., Li, H., Gilitschenski, I., Ivanovic, B., Pavone, M., Geiger, A., Chitta, K.: Pseudo-simulation for autonomous driving. In: CoRL (2025)
4. Chen, B., Martí Monsó, D., Du, Y., Simchowitz, M., Tedrake, R., Sitzmann, V.: Diffusion forcing: Next-token prediction meets full-sequence diffusion. Advances in Neural Information Processing Systems **37**, 24081–24125 (2024)
5. Chen, L., Wu, P., Chitta, K., Jaeger, B., Geiger, A., Li, H.: End-to-end autonomous driving: Challenges and frontiers. IEEE TPAMI (2024)
6. Chi, C., Xu, Z., Feng, S., Cousineau, E., Du, Y., Burchfiel, B., Tedrake, R., Song, S.: Diffusion policy: Visuomotor policy learning via action diffusion. The International Journal of Robotics Research **44**(10-11), 1684–1704 (2025)
7. Chitta, K., Dauner, D., Geiger, A.: SLEDGE: Synthesizing driving environments with generative models and rule-based traffic. In: ECCV (2024)
8. Chung, H., Kim, J., Mccann, M.T., Klasky, M.L., Ye, J.C.: Diffusion posterior sampling for general noisy inverse problems. arXiv preprint arXiv:2209.14687 (2022)
9. Dauner, D., Hallgarten, M., Li, T., Weng, X., Huang, Z., Yang, Z., Li, H., Gilitschenski, I., Ivanovic, B., Pavone, M., et al.: NAVSIM: Data-driven non-reactive autonomous vehicle simulation and benchmarking. In: NeurIPS Datasets and Benchmarks (2024)
10. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. In: NeurIPS (2021)
11. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., Koltun, V.: CARLA: An open urban driving simulator. In: CoRL (2017)
12. Ettinger, S., Cheng, S., Caine, B., Liu, C., Zhao, H., Pradhan, S., Chai, Y., Sapp, B., Qi, C.R., Zhou, Y., et al.: Large scale interactive motion forecasting for autonomous driving: The waymo open motion dataset. In: ICCV (2021)
13. Feng, L., Li, Q., Peng, Z., Tan, S., Zhou, B.: TrafficGen: Learning to generate diverse and realistic traffic scenarios. arXiv preprint arXiv:2210.06609 (2022)
14. Feng, S., Sun, H., Yan, X., Zhu, H., Zou, Z., Shen, S., Liu, H.X.: Dense reinforcement learning for safety validation of autonomous vehicles. Nature **615**(7953), 620–627 (2023)
15. Fiacco, A.V.: Introduction to Sensitivity and Stability Analysis in Nonlinear Programming. New York: Academic Press (1983)
16. Hanselmann, N., Renz, K., Chitta, K., Bhattacharyya, A., Geiger, A.: King: Generating safety-critical driving scenarios for robust imitation via kinematics gradients. In: ECCV (2022)
17. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: NeurIPS (2020)
18. Holderrieth, P., Erives, E.: An introduction to flow matching and diffusion models. arXiv preprint arXiv:2506.02070 (2025)
19. Hu, Y., Chai, S., Yang, Z., Qian, J., Li, K., Shao, W., Zhang, H., Xu, W., Liu, Q.: Solving motion planning tasks with a scalable generative model. In: ECCV (2024)
20. Hu, Y., Yang, J., Chen, L., Li, K., Sima, C., Zhu, X., Chai, S., Du, S., Lin, T., et al.: Planning-oriented autonomous driving. In: CVPR (2023)
21. Huang, Z., Zhang, Z., Vaidya, A., Chen, Y., Fisac, J.F., Lv, C.: Versatile behavior diffusion for generalized traffic agent simulation. IEEE Transactions on Intelligent Transportation Systems (2026)

22. Huang, Z., Zhang, Z., Vaidya, A., Chen, Y., Lv, C., Fisac, J.F.: Versatile scene-consistent traffic scenario generation as optimization with diffusion. arXiv preprint arXiv:2404.02524 (2024)
23. Huang, Z., Zhou, Z., Cai, T., Zhang, Y., Ma, J.: Mdg: Masked denoising generation for multi-agent behavior modeling in traffic environments. arXiv preprint arXiv:2511.17496 (2025)
24. Jiang, C., Cornman, A., Park, C., Sapp, B., Zhou, Y., Anguelov, D., et al.: MotionDiffuser: Controllable multi-agent motion prediction using diffusion. In: CVPR (2023)
25. Jiang, M., Bai, Y., Cornman, A., Davis, C., Huang, X., Jeon, H., Kulshrestha, S., Lambert, J., Li, S., Zhou, X., et al.: SceneDiffuser: Efficient and controllable driving simulation initialization and rollout. In: NeurIPS (2024)
26. Li, T., Qiu, Y., Wu, Z., Lindström, C., Su, P., Nießner, M., Li, H.: MTGS: Multi-traversal gaussian splatting. arXiv preprint arXiv:2503.12552 (2025)
27. Lin, H., Zhang, Y., Ding, W., Wu, J., Zhao, D.: Model-based policy adaptation for closed-loop end-to-end autonomous driving. In: NeurIPS (2025)
28. Liu, H., Li, T., Yang, H., Chen, L., Wang, C., Guo, K., Tian, H., Li, H., Li, H., Lv, C.: Reinforced refinement with self-aware expansion for end-to-end autonomous driving. IEEE TPAMI (2026)
29. Liu, H.X., Feng, S.: Curse of rarity for autonomous vehicles. Nature communications **15**(1), 4808 (2024)
30. Lopez, P.A., Behrisch, M., Bieker-Walz, L., Erdmann, J., Flötteröd, Y.P., Hilbrich, R., Lücken, L., Rummel, J., Wagner, P., Wießner, E.: Microscopic traffic simulation using sumo. In: ITSC (2018)
31. Ma, E., Zhou, L., Tang, T., Zhang, Z., Han, D., Jiang, J., Zhan, K., Jia, P., Lang, X., Sun, H., et al.: Unleashing generalization of end-to-end autonomous driving with controllable long video generation. arXiv preprint arXiv:2406.01349 (2024)
32. Seff, A., Cera, B., Chen, D., Ng, M., Zhou, A., Nayakanti, N., Refaat, K.S., Al-Rfou, R., Sapp, B.: MotionLM: Multi-agent motion forecasting as language modeling. In: ICCV (2023)
33. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics. In: ICML (2015)
34. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: ICLR (2021)
35. Spica, R., Cristofalo, E., Wang, Z., Montijano, E., Schwager, M.: A real-time game theoretic planner for autonomous two-player drone racing. IEEE Transactions on Robotics **36**(5), 1389–1403 (2020)
36. Tan, S., Lambert, J., Jeon, H., Kulshrestha, S., Bai, Y., Luo, J., Anguelov, D., Tan, M., Jiang, C.M.: SceneDiffuser++: City-scale traffic simulation via a generative world model. In: CVPR (2025)
37. Tian, H., Li, T., Liu, H., Yang, J., Qiu, Y., Li, G., Wang, J., Gao, Y., Zhang, Z., Wang, L., Ye, H., Tan, T., Chen, L., Li, H.: Simscale: Learning to drive via real-world simulation at scale. arXiv preprint arXiv:2511.23369 (2025)
38. Wang, J., Pun, A., Tu, J., Manivasagam, S., Sadat, A., Casas, S., Ren, M., Urtasun, R.: AdvSim: Generating safety-critical scenarios for self-driving vehicles. In: CVPR (2021)
39. Xie, Y., Guo, X., Wang, C., Liu, K., Chen, L.: Advdiffuser: Generating adversarial safety-critical driving scenarios via guided diffusion. In: IROS (2024)
40. Xu, C., Petiushko, A., Zhao, D., Li, B.: DiffScene: Diffusion-based safety-critical scenario generation for autonomous vehicles. In: AAAI (2025)
41. Zhang, L., Peng, Z., Li, Q., Zhou, B.: CAT: Closed-loop adversarial training for safe end-to-end driving. In: CoRL (2023)
42. Zheng, Y., Liang, R., ZHENG, K., Zheng, J., Mao, L., Li, J., Gu, W., Ai, R., Li, S.E., Zhan, X., Liu, J.: Diffusion-based planning for autonomous driving with flexible guidance. In: ICLR (2025)

43. Zhong, Z., Rempe, D., Chen, Y., Ivanovic, B., Cao, Y., Xu, D., Pavone, M., Ray, B.: Language-guided traffic simulation via scene-level diffusion. In: Conference on robot learning. pp. 144–177. PMLR (2023)
44. Zhong, Z., Rempe, D., Xu, D., Chen, Y., Veer, S., Che, T., Ray, B., Pavone, M.: Guided conditional diffusion for controllable traffic simulation. arXiv preprint arXiv:2210.17366 (2022)
45. Zhou, Y., Ye, N., Ljungbergh, W., Li, T., Yang, J., Yang, Z., Zhu, H., Petersson, C., Li, H.: Decoupled diffusion sparks adaptive scene generation. In: ICCV (2025)