DISAGREEMENT-AWARE DISTRIBUTION MATCHING DISTILLATION FOR FEW-STEP DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Distribution Matching Distillation (DMD) enables strong few-step diffusion generation by matching the student distribution to a pretrained teacher. Existing improvements to DMD often augment its supervision or modify its optimization machinery. We ask a complementary question: *can DMD diagnose its own updates using signals already produced during training?* DMD inherently computes the discrepancy between a pretrained teacher and a fake model tracking the student distribution, but conventionally uses it primarily to determine the distribution matching direction. We find that the magnitude of this teacher–fake disagreement provides additional diagnostic information. Across extensive experiments, large disagreement in the high-noise regime consistently coincides with substantially higher-energy DMD updates. This reveals a dual role of disagreement: it determines not only *where* the student should move, but also provides an intrinsic signal of *how aggressively* the update should be applied. Based on this observation, we introduce Disagreement-Aware Distribution Matching Distillation (DA-DMD), which activates only in the high-noise regime and continuously modulates the update strength according to teacher–fake disagreement. DA-DMD reuses quantities already available during DMD training and requires no additional model, evaluator, or inference-time modification. We show that it exactly recovers vanilla DMD outside the selected region, reduces update energy only through selected samples, and exhibits redescending influence under extreme disagreement. Across multiple diffusion backbones, DA-DMD consistently improves few-step generation at the same inference cost. On SD3.5-L, DA-DMD improves HPSv2 from 28.23 to 30.93, while achieving a 64.4% average win rate over DMD2.

1 INTRODUCTION

Diffusion models have become a dominant paradigm for high-fidelity image generation, but their iterative sampling procedure often requires many denoising evaluations at inference time (Ho et al., 2020; Song et al., 2020; Karras et al., 2022; Liu et al., 2026; Shao et al., 2026). This has motivated extensive research on compressing pretrained diffusion models into few-step generators, including progressive distillation, consistency-based approaches, and distribution matching methods (Salimans & Ho, 2022; Song et al., 2023; Luo et al., 2023; Yin et al., 2024b;a). Among them, *Distribution Matching Distillation* (DMD) directly aligns the distribution induced by a few-step student with that of a pretrained diffusion teacher, using a frozen teacher and a fake model that tracks the evolving student distribution (Yin et al., 2024b;a). This distribution-level formulation has enabled strong few-step generation across a range of diffusion backbones.

Recent work has improved DMD through better fake-model optimization, auxiliary objectives, reward-guided optimization, and adaptive teacher mechanisms (Yin et al., 2024a; Luo et al., 2026; Bai et al., 2026; Jiang et al., 2025; Wu et al., 2026a;b; Zhou et al., 2024). We take a complementary perspective: *can DMD make better use of the information it already produces during training?* Every DMD update compares the predictions of a pretrained teacher and a fake model at the same noisy sample. This teacher–fake discrepancy determines the direction of distribution matching, but its magnitude varies substantially across samples and noise levels. We therefore investigate whether this native variation also carries information about the behavior of the resulting update.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

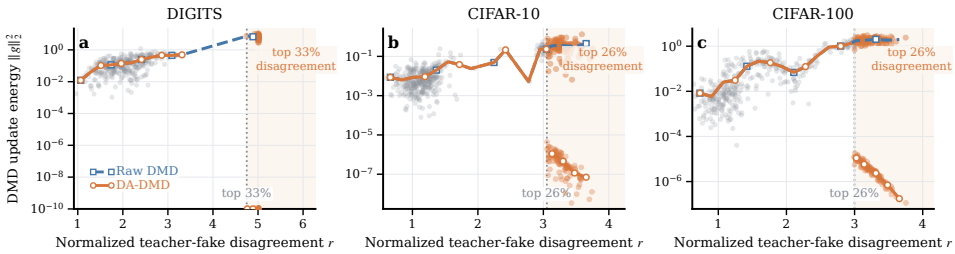

Figure 1: **Teacher–fake disagreement diagnoses DMD update aggressiveness.** Within the predefined high-noise region $\mathcal{H}$, larger normalized disagreement is consistently associated with higher-energy raw DMD updates across DIGITS, CIFAR-10, and CIFAR-100. DA-DMD continuously attenuates these updates according to disagreement. Shaded regions highlight the high-disagreement tail for visualization only and do not correspond to an algorithmic threshold.

Our analysis reveals two complementary patterns. First, on real-world data from DIGITS (Alpaydin & Kaynak, 1998), CIFAR-10 (Krizhevsky et al., 2009), and CIFAR-100 (Krizhevsky et al., 2009), larger normalized teacher–fake disagreement within the high-noise regime consistently coincides with higher-energy raw DMD updates (Fig. 1). Second, the same high-energy behavior persists at scale: during SD3.5-M (Esser et al., 2024) distillation, unusually high-energy raw DMD updates become increasingly concentrated in the high-noise portion of the trajectory (Fig. 2). Importantly, these observations do not imply that high-noise or large-disagreement updates are inherently incorrect. Instead, together they suggest a simple optimization principle: *use noise level to localize where aggressive DMD behavior tends to emerge, and use teacher–fake disagreement to continuously determine how strongly the update should be moderated.*

Motivated by these insights, we introduce **Disagreement-Aware Distribution Matching Distillation (DA-DMD)**. Rather than introducing a new source of guidance, DA-DMD reuses the teacher–fake discrepancy already computed by standard DMD. It activates only in a predefined high-noise region and continuously adjusts the DMD update strength according to normalized teacher–fake disagreement: updates outside this region remain exactly unchanged, while larger disagreement within the region induces stronger attenuation. DA-DMD therefore turns a native DMD signal that originally defines *where* the student should move into an additional control signal for *how conservatively* that move should be applied. Our ablations further show that applying the same correction globally or in lower-noise regions is less effective, supporting the selective high-noise design.

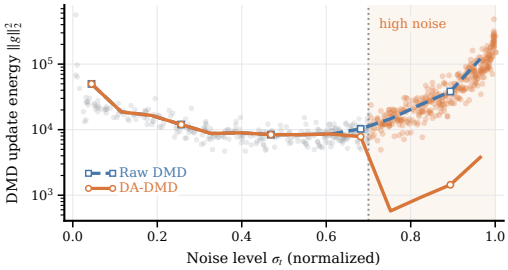

Figure 2: **High-energy DMD updates concentrate at high noise.** During SD3.5-M distillation, raw DMD update energy rises sharply in the high-noise region $\mathcal{H}$, where DA-DMD activates disagreement-aware attenuation.

DA-DMD is theoretically attractive because its robustness is selective rather than global. Outside the active high-noise region, it exactly recovers vanilla DMD, ensuring that ordinary distribution matching supervision is not weakened. Within the active region, disagreement-adaptive weighting reduces update energy only through the modified samples, while its redescending influence prevents extreme teacher–fake disagreement from dominating optimization. Moreover, DA-DMD preserves the ideal fixed point of DMD, modifying the optimization dynamics rather than the target distribution. We conduct extensive experiments across multiple diffusion backbones, including SD3, SD3.5 (Esser et al., 2024), SANA (Xie et al., 2024), and SANA-1.5 (Xie et al., 2025), and evaluate on four complementary benchmarks: DrawBench (Saharia et al., 2022), PartiPrompts (Yu et al., 2022), GenEval2 (Kamath et al., 2025), and TIIF-Bench (Wei et al., 2025). Across model families and evaluation suites, DA-DMD consistently improves the corresponding DMD baselines, with particularly clear gains on challenging compositional and fine-grained generation tasks. These improvements are achieved without changing the inference procedure or sampling cost, while qualitative comparisons further demonstrate better visual fidelity. Our contributions are threefold:

1. We uncover an overlooked role of teacher–fake disagreement, which contains previously underused diagnostic information: in the high-noise regime, larger disagreement is consistently associated with higher-energy DMD updates.
2. We introduce Disagreement-Aware Distribution Matching Distillation (DA-DMD), which reuses this native signal to continuously modulate DMD updates in a predefined high-noise region while exactly preserving vanilla DMD elsewhere.
3. We provide a theoretical analysis of disagreement-aware attenuation. Under standard smoothness and alignment assumptions, we show that adaptive attenuation yields a tighter one-step optimization bound when DMD update energy is sufficiently large relative to its descent signal, providing a principled explanation for selective intervention in the high-noise regime. We further establish exact preservation outside the active region and redescending influence under extreme disagreement, and extensively validate DA-DMD across multiple diffusion backbones and benchmarks.

2 METHOD

Our goal is to make better use of a signal that DMD already computes: the teacher–fake discrepancy. We first revisit how this discrepancy forms the standard DMD update. We then derive an adaptive reweighting from a simple teacher–fake mixture construction and turn it into a practical disagreement-aware update rule. Finally, we provide a theoretical analysis explaining when and why selective attenuation can be preferable to vanilla DMD.

2.1 DISTRIBUTION MATCHING DISTILLATION

Distribution Matching Distillation (DMD) (Yin et al., 2024b;a) trains a few-step generator by matching the noisy marginal distribution induced by the student to that of a pretrained diffusion teacher. Let G_θ denote the student generator and T the frozen teacher. Given condition c and initial noise z, the student generates $x_0^\theta = G_\theta(z, c)$, which is perturbed according to

$$x_t = \alpha_t x_0^\theta + \sigma_t \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I). \tag{1}$$

We denote the resulting student marginal by $q_{\theta,t}(x_t \mid c)$ and the corresponding teacher marginal by $p_{T,t}(x_t \mid c)$. DMD minimizes

$$\mathcal{L}_{\mathrm{DMD}}(\theta) = \mathbb{E}_{t,c}\left[\lambda_t D_{\mathrm{KL}}\left(q_{\theta,t}(\cdot \mid c) \,\|\, p_{T,t}(\cdot \mid c)\right)\right], \tag{2}$$

where λ_t is a timestep-dependent weighting coefficient.

The score of the evolving student marginal is not directly available, so DMD maintains a fake model F_ϕ trained on student-generated samples to track the current student distribution. Let $p_{F,t}$ denote the noisy marginal represented by the fake model, and define $s_T = \nabla_{x_t} \log p_{T,t}(x_t \mid c)$ and $s_F = \nabla_{x_t} \log p_{F,t}(x_t \mid c)$ as the teacher and fake scores, respectively. Absorbing timestep-dependent coefficients into w_t, the sample-wise DMD update can be written as

$$g_{\mathrm{DMD}} = w_t J_\theta^\top \left(s_F - s_T\right), \tag{3}$$

where $J_\theta = \partial x_0^\theta / \partial \theta$.

To connect Eq. 3 with the clean predictions directly produced by diffusion models, we use the Gaussian perturbation identity

$$s_t(x_t, t, c) = \frac{\alpha_t \mathbb{E}[x_0 \mid x_t, c] - x_t}{\sigma_t^2}. \tag{4}$$

For an x_0-prediction model trained with squared-error denoising, $\hat{x}_0(x_t, t, c) \approx \mathbb{E}[x_0 \mid x_t, c]$, and therefore

$$s_t(x_t, t, c) \approx \frac{\alpha_t \hat{x}_0(x_t, t, c) - x_t}{\sigma_t^2}. \tag{5}$$

A complete derivation of Eq. 4 from Fisher's identity is provided in Appendix C.1.

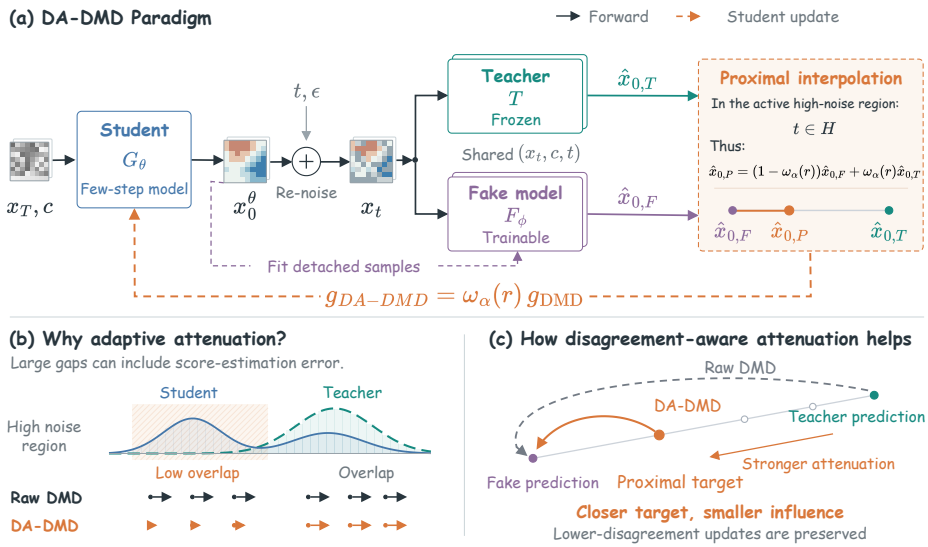

Figure 3: **Overview of DA-DMD.** (a) In the active high-noise region, DA-DMD replaces the raw teacher-directed correction with a disagreement-aware proximal interpolation between the fake and teacher predictions, yielding an attenuated DMD update with multiplier $\omega_\alpha(r)$. (b) Large teacher–fake disagreement can induce disproportionately large raw DMD corrections. (c) DA-DMD shortens these corrections toward an adaptive proximal target, reducing the influence of high-disagreement updates while preserving lower-disagreement updates.

Since the teacher and fake model are evaluated at the same noisy sample x_t, the shared term $-x_t/\sigma_t^2$ cancels, giving

$$s_F - s_T \approx \frac{\alpha_t}{\sigma_t^2}\left(\hat{x}_{0,F} - \hat{x}_{0,T}\right). \tag{6}$$

Substituting Eq. 6 into Eq. 3 yields

$$g_{\mathrm{DMD}} \approx \gamma_t J_\theta^\top \left(\hat{x}_{0,F} - \hat{x}_{0,T}\right), \tag{7}$$

where $\gamma_t = w_t \alpha_t / \sigma_t^2$ collects the timestep- and objective-dependent scalar factors.

Eq. 7 exposes the central signal driving DMD: the discrepancy between the teacher and fake predictions. Its direction determines the orientation of the distribution-matching correction, while its magnitude directly affects the resulting update strength. Importantly, this discrepancy varies substantially across samples and noise levels, as observed in Sec. 1. This raises a complementary question: *can the teacher–fake discrepancy be used not only to construct the DMD correction, but also to determine how strongly that correction should be applied?*

2.2 From Distribution Mixture to Disagreement-Aware DMD

From distribution mixture to adaptive DMD. To derive an adaptive DMD correction, we begin with a simple auxiliary mixture of the teacher and fake distributions. For clarity, we omit the conditioning variable c. At noise level t, define

$$p_{\alpha,t}(x_t) = \alpha p_{T,t}(x_t) + (1-\alpha) p_{F,t}(x_t), \qquad \alpha \in (0,1). \tag{8}$$

Although α is fixed at the distribution level, the corresponding score mixture is sample dependent. Let $\Delta_t(x_t) = \log \frac{p_{T,t}(x_t)}{p_{F,t}(x_t)}$, and define the generic mixture weighting function

$$\mathcal{W}_\alpha(u) = \frac{\alpha e^u}{\alpha e^u + 1 - \alpha}. \tag{9}$$

Using $p_{\alpha,t}(x_t) = p_{F,t}(x_t)[\alpha e^{\Delta_t(x_t)} + 1 - \alpha]$ and $\nabla_{x_t}\Delta_t = s_T - s_F$, the score of the auxiliary mixture becomes

$$s_\alpha = \mathcal{W}_\alpha(\Delta_t) s_T + [1 - \mathcal{W}_\alpha(\Delta_t)]\, s_F. \tag{10}$$

Therefore, $s_F - s_\alpha = \mathcal{W}_\alpha(\Delta_t)(s_F - s_T)$, which induces the adaptive DMD correction

$$g_\alpha = \mathcal{W}_\alpha(\Delta_t) g_{\mathrm{DMD}}. \tag{11}$$

This sample adaptivity is specific to the arithmetic mixture. By contrast, the geometric interpolation $p^{\mathrm{geo}}_{\alpha,t} \propto p^{\alpha}_{T,t} p^{1-\alpha}_{F,t}$ yields $s^{\mathrm{geo}}_\alpha = \alpha s_T + (1-\alpha) s_F$, and therefore only a uniform rescaling of the DMD correction by α. We therefore adopt the arithmetic mixture to obtain sample-dependent weighting; a detailed comparison is provided in Appendix C.6.

From density ratio to native disagreement. The exact adaptive weight in Eq. 10 depends on the sample-wise log-density ratio $\Delta_t(x_t)$, which is unavailable in standard diffusion models. We therefore instantiate the same weighting principle using a tractable signal already produced during DMD training. From Eq. 6, the teacher–fake clean-prediction discrepancy is directly related to the corresponding score discrepancy. We define its noise-normalized magnitude as

$$r^2 = \mathrm{mean}\left[\left(\frac{\hat{x}_{0,F} - \hat{x}_{0,T}}{\sigma_t}\right)^2\right], \tag{12}$$

where the mean is taken over the non-batch latent dimensions. By Eq. 6, $(\hat{x}_{0,F} - \hat{x}_{0,T})/\sigma_t \approx (\sigma_t/\alpha_t)(s_F - s_T)$, so r measures the magnitude of the same native teacher–fake discrepancy that drives the DMD correction, after accounting for its noise-dependent scale.

To convert this discrepancy into an adaptive weight, we use the monotone quadratic-energy surrogate $\widetilde{\Delta}(r) = -r^2/2$ for the inaccessible sample-wise density evidence. Importantly, $\widetilde{\Delta}(r)$ is not intended to estimate the true log-density ratio $\Delta_t(x_t)$; it provides a conservative, disagreement-dependent surrogate within the same mixture-weighting form. The practical DA-DMD weight is therefore

$$\omega_\alpha(r) := \mathcal{W}_\alpha\left(-\frac{r^2}{2}\right) = \frac{\alpha e^{-r^2/2}}{\alpha e^{-r^2/2} + 1 - \alpha}. \tag{13}$$

The weight decreases monotonically with disagreement, satisfying $\omega_\alpha(0) = \alpha$ and $\omega_\alpha(r) \to 0$ as $r \to \infty$. Hence, the same coefficient α that defines the auxiliary mixture also determines the maximum retained DMD update strength in the practical weighting rule, while disagreement introduces additional sample-adaptive attenuation.

Selective high-noise intervention. The analysis in Sec. 1 shows that unusually high-energy DMD updates concentrate in the high-noise regime. We therefore restrict disagreement-aware attenuation to a predefined high-noise region $\mathcal{H}$. Let $\chi = \mathbf{1}[t \in \mathcal{H}]$. The final DA-DMD update is

$$g_{\mathrm{DA-DMD}} = \mathrm{sg}\left[1 - \chi + \chi \omega_\alpha(r)\right] g_{\mathrm{DMD}}, \tag{14}$$

where $\mathrm{sg}[\cdot]$ denotes stop-gradient. Equivalently,

$$g_{\mathrm{DA-DMD}} = \begin{cases} g_{\mathrm{DMD}}, & t \notin \mathcal{H}, \\ \omega_\alpha(r) g_{\mathrm{DMD}}, & t \in \mathcal{H}. \end{cases}$$

Eq. 14 separates the roles of noise level and disagreement: the high-noise gate determines *where* DA-DMD intervenes, while normalized teacher–fake disagreement determines *how strongly* the update is attenuated. Appendix Fig. 7 visualizes this weighting behavior, including the exact preservation outside $\mathcal{H}$ and the progressively stronger attenuation within the active region.

The resulting effect on optimization is illustrated in Appendix Fig. 8, and Fig. 9 and 10. In the controlled example, high-disagreement samples are associated with larger raw DMD updates, whereas DA-DMD reduces their influence while retaining the underlying DMD correction direction. This provides an intuitive view of how disagreement-aware attenuation can prevent a small subset of aggressive updates from disproportionately shaping the optimization trajectory.

All quantities required by DA-DMD are already available during standard DMD training. In all experiments, both our DMD2 baseline and DA-DMD use only the score-based distribution-matching branch, without the real-data adversarial objective. DA-DMD therefore modifies only this DMD update through disagreement-aware reweighting, while the student rollout, teacher query, fake-model optimization, and inference procedure remain unchanged.

Algorithm 1 DA-DMD

Require: Teacher T, student G_θ, fake model F_ϕ; weighting coefficient $\alpha \in (0,1)$; noise schedule $\{\sigma_t\}$; high-noise region $\mathcal{H}$

for each training iteration **do**
  Obtain x_0^θ, x_t, $\hat{x}_{0,T}$, $\hat{x}_{0,F}$, and g_{DMD} as in DMD2 ▷ standard DMD
  $r^2 \leftarrow \mathrm{mean}\left[\left(\frac{\hat{x}_{0,F}-\hat{x}_{0,T}}{\sigma_t}\right)^2\right]$ ▷ measure normalized disagreement
  $\chi \leftarrow \mathbf{1}[t \in \mathcal{H}]$ ▷ activate in high-noise region
  $\omega_\alpha(r) \leftarrow \frac{\alpha e^{-r^2/2}}{\alpha e^{-r^2/2} + 1 - \alpha}$ ▷ compute adaptive retention
  $g_{\mathrm{DA-DMD}} \leftarrow \mathrm{sg}[1 - \chi + \chi\omega_\alpha(r)]\, g_{\mathrm{DMD}}$ ▷ selectively reweight DMD
  Update G_θ with $g_{\mathrm{DA-DMD}}$; update F_ϕ as in DMD2
end for

2.3 DA-DMD as Selective Adaptive Step Control

The mixture construction in Sec. 2.2 explains how the adaptive weighting rule arises. We now turn to a complementary question: when is attenuating a DMD update actually preferable to applying it at full strength? Importantly, DA-DMD does not interpret high disagreement as evidence of an incorrect update. Instead, our analysis considers a subtler regime in which the update remains aligned with a useful descent direction but its magnitude becomes overly aggressive.

Let $g = g_{\mathrm{DMD}}$, and consider a generic attenuated step $\theta^+ = \theta - \eta a g$, where $a \in [0,1]$. Assume that the DMD objective $\mathcal{L}(\theta)$ is β-smooth and that the conditional DMD update has positive descent alignment. Conditioned on noise level t and disagreement r, define $A(t,r) = \mathbb{E}[\langle \nabla\mathcal{L}(\theta), g\rangle \mid t, r] > 0$ as the useful descent alignment and $B(t,r) = \mathbb{E}[\|g\|^2 \mid t, r] > 0$ as the update energy. The standard smoothness inequality gives

$$\mathbb{E}\left[\mathcal{L}(\theta^+) \mid t, r\right] \le \mathcal{L}(\theta) - \eta a A(t,r) + \frac{\beta\eta^2 a^2}{2} B(t,r). \tag{15}$$

Minimizing the right-hand side over $a \in [0,1]$ yields

$$a^\star(t,r) = \min\left\{1, \frac{A(t,r)}{\beta\eta B(t,r)}\right\}. \tag{16}$$

Eq. 16 exposes the key optimization principle behind DA-DMD: *the preferred update strength decreases when update energy becomes large relative to useful descent alignment*. Thus, a DMD update can remain positively aligned with the descent direction while still being too aggressive to apply at full strength. In other words, a useful update direction does not necessarily imply an appropriate update magnitude. More explicitly, for any fixed attenuation factor $a \in (0,1)$, the corresponding conditional one-step bound is tighter than that of vanilla DMD whenever $B(t,r)/A(t,r) > 2/[\beta\eta(1+a)]$. The complete derivation is provided in Appendix C.2.

The analysis further connects the empirical disagreement signal to adaptive step control. Within the active high-noise region $\mathcal{H}$, suppose that the conditional energy-to-alignment ratio $B(t,r)/A(t,r)$ is nondecreasing with disagreement r. Eq. 16 then implies

$$r \uparrow \quad \Longrightarrow \quad a^\star(t,r) \downarrow .$$

Such behavior is consistent with the practical DA-DMD weight $\omega_\alpha(r)$ in Eq. 13, which assigns stronger attenuation to larger disagreement. Under this assumption, teacher–fake disagreement therefore serves as an intrinsic control signal that follows the same monotonic step-control principle suggested by the smoothness analysis.

DA-DMD further localizes this robustification to the region where aggressive updates empirically concentrate. Outside the active region $\mathcal{H}$, vanilla DMD is recovered exactly, while any induced reduction in update energy is confined to the high-noise regime. In addition, the disagreement-dependent influence is redescending:

$$t \notin \mathcal{H} \implies g_{\mathrm{DA-DMD}} = g_{\mathrm{DMD}}, \qquad \lim_{r\to\infty} r\,\omega_\alpha(r) = 0.$$

Table 1: **Quantitative comparison on DrawBench.** DA-DMD provides broad improvements over DMD2 across multiple diffusion backbones and evaluation metrics.

Backbone	Method	PickScore↑	ImageReward↑	HPSv2↑	AES↑	VQ-R1↑	Avg. WR↑
SANA	DMD2 (Yin et al., 2024a)	22.53	95.05	29.67	5.820	4.580	-
	DA-DMD (Ours)	**22.56** +0.03	**101.58** +6.53	**29.87** +0.20	**5.945** +0.125	**4.643** +0.063	**56.0%**
SANA-1.5	DMD2 (Yin et al., 2024a)	22.53	98.57	29.06	5.858	4.525	-
	DA-DMD (Ours)	**22.58** +0.05	**102.09** +3.52	**29.35** +0.29	**5.916** +0.058	**4.549** +0.024	**57.0%**
SD3-M	DMD2 (Yin et al., 2024a)	22.33	98.56	30.00	5.463	4.534	-
	Hyper-SD3 (Ren et al., 2024)	21.99	67.67	26.30	5.305	4.292	28.8%
	SiD-DiT (Zhou et al., 2024)	22.46	91.31	28.01	5.287	4.599	39.8%
	SiD-Adv (Zhou et al., 2024)	22.53	90.67	29.62	**5.560**	4.624	50.2%
	TDM (Luo et al., 2025)	22.49	101.13	29.81	5.474	**4.693**	52.2%
	DA-DMD (Ours)	**22.56** +0.23	**110.72** +12.16	**30.39** +0.39	5.556 +0.093	4.649 +0.115	**61.2%**
SD3.5-M	DMD2 (Yin et al., 2024a)	22.49	97.82	28.70	5.366	4.647	-
	SiD-DiT (Zhou et al., 2024)	**22.57**	91.15	28.92	5.463	4.682	48.0%
	SiD-Adv (Zhou et al., 2024)	22.50	93.75	28.97	**5.574**	4.558	48.8%
	DP-DMD (Wu et al., 2026a)	22.32	91.37	28.08	5.299	4.553	39.4%
	SenseFlow (Ge et al., 2025)	22.42	89.37	29.43	5.564	4.648	56.8%
	DA-DMD (Ours)	22.51 +0.02	**104.20** +6.38	**29.57** +0.87	5.474 +0.108	**4.700** +0.053	**59.5%**
SD3.5-L	DMD2 (Yin et al., 2024a)	22.56	94.33	28.23	5.378	4.550	-
	SiD-DiT (Zhou et al., 2024)	22.67	101.77	29.35	5.543	4.566	53.8%
	SiD-Adv (Zhou et al., 2024)	22.69	112.92	30.52	**5.743**	4.638	**65.6%**
	SD3.5-L Turbo (Sauer et al., 2024)	22.65	91.07	28.47	5.631	4.608	54.4%
	SenseFlow (Ge et al., 2025)	22.70	112.52	30.18	5.421	**4.652**	62.2%
	DA-DMD (Ours)	**22.73** +0.17	**113.88** +19.55	**30.93** +2.70	5.506 +0.128	4.587 +0.037	64.4%

Consequently, DA-DMD moderates aggressive high-noise updates without globally weakening the original distribution-matching signal, while progressively reducing the influence of extreme teacher–fake disagreement. Complete proofs of these properties, together with complementary bias–variance and proximal interpretations, are provided in Appendix C.3.

3 Experiments

In this section, we conduct extensive experiments to validate DA-DMD. Following the framework developed in Secs. 2.2 and 2.3, these experiments primarily serve to verify the effectiveness of disagreement-aware update control and the rationality of our selective high-noise design across multiple diffusion backbones and text-to-image benchmarks.

3.1 Experimental Setup

We evaluate DA-DMD on 5 text-to-image diffusion backbones, including SANA (Xie et al., 2024), SANA-1.5 (Xie et al., 2025), SD3-M, SD3.5-M, and SD3.5-L (Esser et al., 2024). Our evaluation covers 4 complementary benchmarks: DrawBench (Saharia et al., 2022), PartiPrompts (Yu et al., 2022), GenEval2 (Kamath et al., 2025), and TIIF-Bench (Wei et al., 2025), spanning general text-to-image generation, compositional fidelity, and fine-grained instruction following. We report PickScore (Kirstain et al., 2023), ImageReward (Xu et al., 2023), HPSv2 (Wu et al., 2023), AES (Schuhmann, 2022), VisualQuality-R1 (Wu et al., 2026c), and average pairwise win rate as complementary measures of preference alignment and visual quality. DMD2 (Yin et al., 2024a) serves as our primary baseline, together with representative few-step distillation and acceleration methods where applicable. Detailed benchmark descriptions, evaluation protocols, model configurations, and implementation settings are provided in the Appendix Sec. A.

3.2 Main Results

Quantitative comparison. Tab. 1 and Appendix Tab. 3 report the main results on DrawBench and PartiPrompts. Across all 5 backbones, DA-DMD consistently improves the corresponding DMD2

baseline on both benchmarks, demonstrating that the benefit of disagreement-aware update control generalizes across different model families and scales. The gains are particularly pronounced on the SD3/SD3.5 family. On DrawBench, DA-DMD achieves average pairwise win rates of 61.2%, 59.5%, and 64.4% on SD3-M, SD3.5-M, and SD3.5-L, respectively, while improving multiple preference- and quality-oriented metrics. Similar trends are observed on PartiPrompts, where DA-DMD reaches 61.2% and 61.4% average win rates on SD3.5-M and SD3.5-L. Compared with representative few-step distillation and acceleration baselines, DA-DMD remains competitive or superior across most evaluation dimensions while preserving the original inference procedure and sampling cost. Additional results on GenEval2 and TIIF-Bench in the Appendix Tab. 4, 5, 6, 7 further show that these improvements extend to compositional generation and fine-grained instruction following.

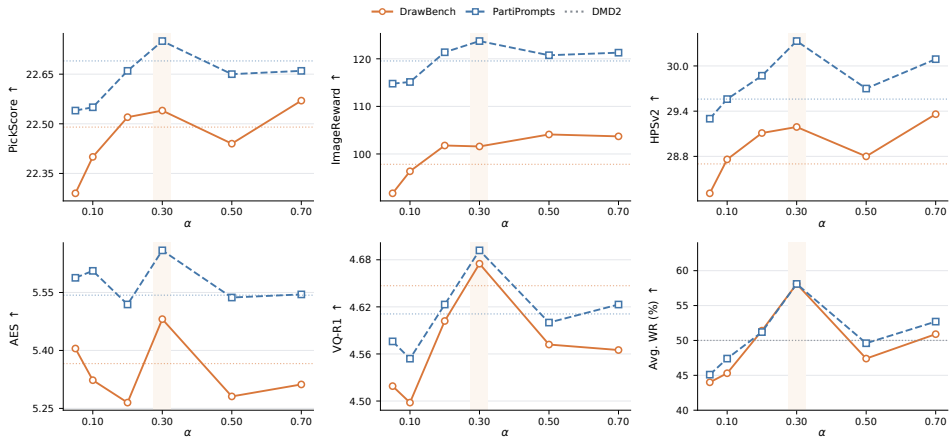

Figure 4: **Sensitivity to the weighting coefficient α on SD3.5-M.** Moderate weighting, particularly $\alpha = 0.3$, provides the best overall trade-off across metrics on DrawBench and PartiPrompts.

Diagnosing DMD updates. Beyond generation quality, we examine whether teacher–fake disagreement provides a meaningful diagnostic of DMD optimization. Across DIGITS, CIFAR-10, and CIFAR-100, high-disagreement samples produce larger raw DMD updates and can disproportionately steer the optimization trajectory, whereas DA-DMD attenuates these corrections and recovers target-mode occupancy more faithfully (Figs. 8, 9, and 10). Consistently, larger normalized disagreement within the high-noise regime coincides with higher-energy raw DMD updates across all three datasets (Fig. 1). At model scale, SD3.5-M further shows that high-energy updates concentrate in the high-noise portion of the trajectory (Fig. 2). Together, these observations support the design of DA-DMD: noise level determines *where* to intervene, while disagreement determines *how strongly*.

3.3 Ablation Study

Where should DA-DMD intervene? We first investigate where disagreement-aware attenuation should be applied. As shown in Fig. 6, restricting DA-DMD to the high-noise region provides the strongest overall trade-off across DrawBench and PartiPrompts, with particularly clear gains in average pairwise win rate and most preference- and quality-oriented metrics. In contrast, global, middle-noise, and low-noise intervention is generally less effective and can underperform vanilla DMD on several metrics. These results indicate that disagreement-aware weighting is not uniformly beneficial across the diffusion trajectory, supporting our selective high-noise design.

Sensitivity to the weighting coefficient. We next study the effect of the weighting coefficient α, which controls the maximum retained update strength in the active region through $\omega_\alpha(0) = \alpha$. Fig. 4 shows that moderate weighting consistently provides the best overall trade-off across evaluation metrics, with $\alpha = 0.3$ performing particularly well on both DrawBench and PartiPrompts. Smaller values tend to over-attenuate useful updates, whereas larger values provide weaker correction. Overall, DA-DMD exhibits a clear and interpretable sensitivity to α, favoring moderate intervention over either extreme.

3.4 Qualitative Results

Fig. 5 and Appendix Fig. 11 compare DA-DMD with representative few-step distillation and acceleration baselines on SD3.5-M. Across challenging prompts involving counting, text rendering,

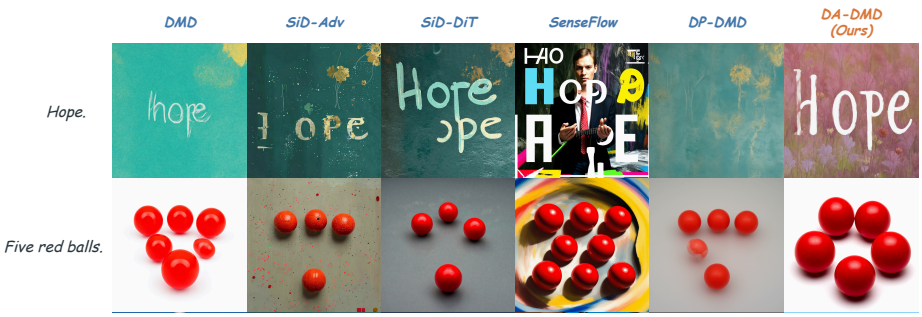

Figure 5: Qualitative comparison with representative distillation baselines on SD3.5-M.

attribute binding, and object identity, DA-DMD produces generations that more faithfully satisfy the specified visual constraints while maintaining strong perceptual quality. These qualitative improvements are consistent with the quantitative gains reported in Sec. 3.2.

We further compare DA-DMD directly against DMD2 across all evaluated backbones. As shown in Appendix Fig. 12, 13, 14, 15, and 16, DA-DMD consistently improves fine-grained prompt adherence across diverse compositional challenges, including object counting, spatial relations, relative attributes, color control, and structured multi-object generation. Together, these examples suggest that disagreement-aware update control improves not only automatic evaluation scores but also visually observable compositional fidelity.

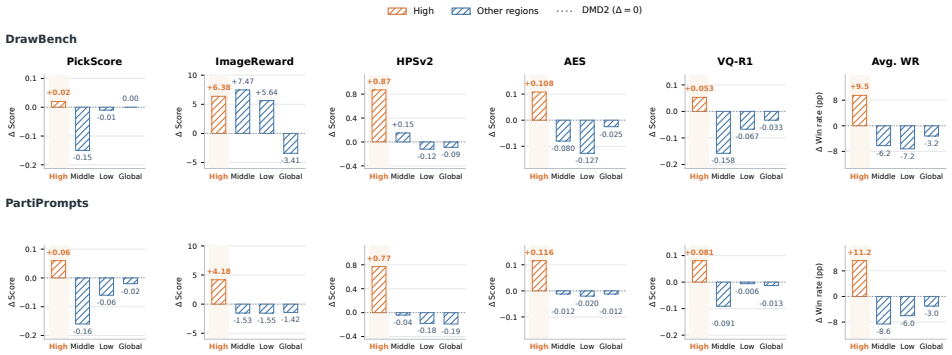

Figure 6: **Effect of the intervention region on SD3.5-M.** High-noise intervention provides the strongest overall trade-off across DrawBench and PartiPrompts, while global, middle-noise, and low-noise attenuation are generally less effective.

Limitations. DA-DMD currently relies on a predefined high-noise intervention region and a hand-designed disagreement-to-weight mapping. While these choices are effective across the evaluated settings, the preferred intervention strength can vary across model families. A promising direction for future work is therefore to make both components more adaptive, allowing the model to determine when intervention is beneficial and how strongly the DMD correction should be modulated.

Conclusion. DA-DMD reuses the native teacher–fake discrepancy in DMD as an intrinsic signal for adaptive update control. Our key observation is that, beyond forming the distribution-matching correction itself, the magnitude of this discrepancy also provides useful information about the appropriate update strength, particularly in the high-noise regime where aggressive corrections tend to concentrate. Based on this observation, DA-DMD selectively attenuates high-noise updates according to disagreement while exactly preserving vanilla DMD outside the active region. Our analysis interprets this mechanism as adaptive step control, providing a principled explanation for why a directionally useful DMD correction can still benefit from a reduced step magnitude. Extensive experiments across multiple diffusion backbones and text-to-image benchmarks demonstrate broad improvements over DMD2 without introducing additional supervision or inference-time overhead. More broadly, our results suggest that native signals already present in distribution matching can be leveraged not only to construct the correction, but also to control how strongly it is applied.

AI USE STATEMENT

Generative AI tools were used to assist with research code implementation, manuscript organization, language polishing, and the presentation of mathematical derivations. All AI-assisted content was independently reviewed and verified by the authors, including manual inspection of generated code, mathematical arguments, experimental analyses, and manuscript revisions. The final research decisions, experimental design, technical claims, and scientific conclusions were made and validated by the authors, who take full responsibility for the content of this work.

ETHICS STATEMENT

This work studies efficient distillation of text-to-image diffusion models and aims to reduce the computational cost of high-quality image generation. As with other generative modeling techniques, accelerated generation may inherit broader risks associated with image synthesis, including misuse for deceptive or inappropriate content. DA-DMD does not introduce new data collection, user profiling, or additional inference-time models, and is intended for responsible research and legitimate generative applications.

REPRODUCIBILITY STATEMENT

We provide the formulation and training procedure of DA-DMD, including the disagreement measure, adaptive weighting rule, high-noise intervention strategy, and complete training pseudocode. We evaluate the method across multiple diffusion backbones and benchmarks, with experimental settings, evaluation protocols, ablations, and additional theoretical derivations provided in the Appendix. These details are intended to support reproducibility and facilitate comparison with existing DMD-based distillation methods.

REFERENCES

E. Alpaydin and C. Kaynak. Optical Recognition of Handwritten Digits. UCI Machine Learning Repository, 1998.

Lichen Bai, Zikai Zhou, Shitong Shao, Wenliang Zhong, Shuo Yang, Shuo Chen, Bojun Chen, and Zeke Xie. Optimizing few-step generation with adaptive matching distillation. *arXiv preprint arXiv:2602.07345*, 2026.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, et al. Scaling rectified flow Transformers for high-resolution image synthesis. In *International Conference on Machine Learning*, 2024.

Xingtong Ge, Xin Zhang, Tongda Xu, Yi Zhang, Xinjie Zhang, Yan Wang, and Jun Zhang. Senseflow: Scaling distribution matching for flow-based text-to-image distillation. *arXiv preprint arXiv:2506.00523*, 2025.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 2020.

Dengyang Jiang, Dongyang Liu, Zanyi Wang, Qilong Wu, Liuzhuozheng Li, Hengzhuang Li, Xin Jin, David Liu, Changsheng Lu, Zhen Li, et al. Distribution matching distillation meets reinforcement learning. *arXiv preprint arXiv:2511.13649*, 2025.

Amita Kamath, Kai-Wei Chang, Ranjay Krishna, Luke Zettlemoyer, Yushi Hu, and Marjan Ghazvininejad. Geneval 2: Addressing benchmark drift in text-to-image evaluation. *arXiv preprint arXiv:2512.16853*, 2025.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 2022.

Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. *Advances in neural information processing systems*, 2023.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

Buhua Liu, Shitong Shao, Bao Li, Lichen Bai, Zhiqiang Xu, Haoyi Xiong, James T. Kwok, Sumi Helal, and Zeke Xie. Alignment of diffusion models: Fundamentals, challenges, and future. 2026.

Dongyang Liu, Peng Gao, David Liu, Ruoyi Du, Zhen Li, Qilong Wu, Xin Jin, Sihan Cao, Shifeng Zhang, Hongsheng Li, et al. Decoupled DMD: CFG augmentation as the spear, distribution matching as the shield. *arXiv preprint arXiv:2511.22677*, 2025.

Simian Luo, Yiqin Tan, Longbo Huang, Jian Li, and Hang Zhao. Latent consistency models: Synthesizing high-resolution images with few-step inference. *arXiv preprint arXiv:2310.04378*, 2023.

Yihong Luo, Tianyang Hu, Jiacheng Sun, Yujun Cai, and Jing Tang. Learning few-step diffusion models by trajectory distribution matching. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, 2025.

Yihong Luo, Tianyang Hu, Weijian Luo, and Jing Tang. TDM-R1: Reinforcing few-step diffusion models with non-differentiable reward. *arXiv preprint arXiv:2603.07700*, 2026.

Yuxi Ren, Xin Xia, Yanzuo Lu, Jiacheng Zhang, Jie Wu, Pan Xie, Xing Wang, and Xuefeng Xiao. Hyper-SD: Trajectory segmented consistency model for efficient image synthesis. *Advances in neural information processing systems*, 2024.

Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. Photorealistic text-to-image diffusion models with deep language understanding. *Advances in neural information processing systems*, 2022.

Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. *arXiv preprint arXiv:2202.00512*, 2022.

Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion distillation. In *Computer Vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LXXXVI*, 2024.

Christoph Schuhmann. Improved aesthetic predictor. https://github.com/christophschuhmann/improved-aesthetic-predictor, 2022. GitHub repository.

Shitong Shao, Lichen Bai, Pengfei Wan, James Kwok, and Zeke Xie. Efficient video diffusion models: Advancements and challenges, 2026. URL https://arxiv.org/abs/2604.15911.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. *arXiv preprint arXiv:2010.02502*, 2020.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International Conference on Machine Learning*, 2023.

Xinyu Wei, Jinrui Zhang, Zeqing Wang, Hongyang Wei, Zhen Guo, and Lei Zhang. TIIF-Bench: How does your T2I model follow your instructions? *arXiv preprint arXiv:2506.02161*, 2025.

Tianhe Wu, Ruibin Li, Lei Zhang, and Kede Ma. Diversity-preserved distribution matching distillation for fast visual synthesis. *arXiv preprint arXiv:2602.03139*, 2026a.

Tianhe Wu, Kun Yan, Zikai Zhou, Lihan Jiang, Jiahao Li, Jie Zhang, Kaiyuan Gao, Ningyuan Tang, Shengming Yin, Xiaoyue Chen, Xiao Xu, Yilei Chen, Yuxiang Chen, Yan Shu, Yixian Xu, Yanran Zhang, Zihao Liu, Zhendong Wang, Zekai Zhang, Deqing Li, Liang Peng, Yi Wang, Jingren Zhou, and Chenfei Wu. Qwen-Image-Flash: Beyond objective design, 2026b. URL https://arxiv.org/abs/2606.03746.

Tianhe Wu, Jian Zou, Jie Liang, Lei Zhang, and Kede Ma. Visualquality-r1: Reasoning-induced image quality assessment via reinforcement learning to rank. *Advances in Neural Information Processing Systems*, 2026c.

Xiaoshi Wu, Yiming Hao, Keqiang Sun, Yixiong Chen, Feng Zhu, Rui Zhao, and Hongsheng Li. Human Preference Score v2: A solid benchmark for evaluating human preferences of text-to-image synthesis. *arXiv preprint arXiv:2306.09341*, 2023.

Enze Xie, Junsong Chen, Junyu Chen, Han Cai, Haotian Tang, Yujun Lin, Zhekai Zhang, Muyang Li, Ligeng Zhu, Yao Lu, et al. Sana: Efficient high-resolution image synthesis with linear diffusion transformers. *arXiv preprint arXiv:2410.10629*, 2024.

Enze Xie, Junsong Chen, Yuyang Zhao, Jincheng Yu, Ligeng Zhu, Chengyue Wu, Yujun Lin, Zhekai Zhang, Muyang Li, Junyu Chen, et al. Sana 1.5: Efficient scaling of training-time and inference-time compute in linear diffusion transformer. *arXiv preprint arXiv:2501.18427*, 2025.

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. Imagereward: Learning and evaluating human preferences for text-to-image generation. *Advances in Neural Information Processing Systems*, 2023.

Tianwei Yin, Michaël Gharbi, Taesung Park, Richard Zhang, Eli Shechtman, Fredo Durand, and Bill Freeman. Improved distribution matching distillation for fast image synthesis. In *Advances in Neural Information Processing Systems*, pp. 47455–47487, 2024a.

Tianwei Yin, Michaël Gharbi, Richard Zhang, Eli Shechtman, Fredo Durand, William T Freeman, and Taesung Park. One-step diffusion with distribution matching distillation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6613–6623, 2024b.

Jiahui Yu, Yuanzhong Xu, Jing Yu Koh, Thang Luong, Gunjan Baid, Zirui Wang, Vijay Vasudevan, Alexander Ku, Yinfei Yang, Burcu Karagol Ayan, et al. Scaling autoregressive models for content-rich text-to-image generation. *arXiv preprint arXiv:2206.10789*, 2022.

Mingyuan Zhou, Huangjie Zheng, Zhendong Wang, Mingzhang Yin, and Hai Huang. Score identity distillation: Exponentially fast distillation of pretrained diffusion models for one-step generation. In *Forty-first International Conference on Machine Learning*, 2024.

A Implementation Details

In this section, we provide an overview of the benchmarks, evaluation metrics, diffusion models and the hyperparameter settings in our main paper.

A.1 Benchmark

DrawBench. We evaluate general text-to-image generation capabilities using DrawBench (Saharia et al., 2022), a challenging benchmark designed to probe a broad range of semantic and compositional skills. DrawBench contains 200 manually curated prompts organized into 11 categories, covering challenging aspects such as color and attribute binding, counting, spatial relations, text rendering, rare concepts, misspellings, and unusual object interactions. It further includes complex and long-form descriptions that require models to correctly compose multiple visual concepts. The diversity and compact size of DrawBench make it particularly suitable for fine-grained evaluation of both prompt alignment and perceptual generation quality.

PartiPrompts. We further evaluate on PartiPrompts (P2) (Yu et al., 2022), a large-scale and diverse benchmark introduced for evaluating content-rich text-to-image generation. PartiPrompts consists of 1,632 English prompts spanning 12 semantic categories and 11 challenge aspects. The semantic categories cover a broad range of visual concepts, including people, animals, scenes, artifacts, abstract concepts, and world knowledge, while the challenge aspects range from basic generation and simple details to fine-grained attributes, quantity, spatial positioning, linguistic structures, writing and symbols, and complex compositions. This structured organization enables evaluation across both prompt content and generation difficulty, making PartiPrompts particularly useful for assessing generalization to diverse and compositionally demanding instructions.

TIIF-Bench. To evaluate fine-grained instruction-following ability, we adopt TIIF-Bench (Wei et al., 2025), which is specifically designed to assess how accurately modern text-to-image models interpret and execute complex user instructions. TIIF-Bench contains 5,000 diverse prompts organized across multiple dimensions and three levels of difficulty and complexity. For each instruction, the benchmark provides both a concise version and a rhetorically enriched long-form version with the same underlying semantics, allowing systematic evaluation of robustness to prompt length and linguistic complexity. Beyond conventional compositional requirements, TIIF-Bench also covers challenging capabilities such as text rendering, style control, and designer-level generation tasks. These characteristics make it well suited for measuring fine-grained instruction adherence under realistic and complex generation settings.

GenEval 2. We additionally employ GenEval 2 (Kamath et al., 2025), an updated compositional text-to-image benchmark designed to address the increasing saturation and benchmark drift of earlier automatic evaluation suites. GenEval 2 contains 800 prompts with substantially improved coverage of primitive visual concepts and higher degrees of compositionality, including objects, attributes, spatial relations, and counting. Each prompt is decomposed into a set of atomic visual requirements, enabling evaluation not only at the prompt level but also at the level of individual compositional skills. GenEval 2 is evaluated using Soft-TIFA, a VQA-based evaluation framework that aggregates soft judgments over these visual primitives and is designed to better align automatic evaluation with human judgment. The benchmark therefore provides a fine-grained diagnostic of whether a model can simultaneously satisfy multiple visual constraints in increasingly complex prompts.

A.2 Evaluation Metrics

PickScore. PickScore (Kirstain et al., 2023) is a CLIP-based human-preference metric trained on the Pick-a-Pic dataset, which contains pairwise preferences collected from real users of text-to-image systems. It jointly considers the input prompt and generated image to estimate human preference, providing an automatic measure of both semantic alignment and overall generation quality.

ImageReward. ImageReward (Xu et al., 2023) is a general-purpose text-to-image reward model trained on expert-annotated image comparisons. It is designed to capture human preferences over

generated images while accounting for both text–image alignment and perceptual quality, and has been widely adopted for automatic evaluation of text-to-image generation.

HPS v2. Human Preference Score v2 (HPS v2) (Wu et al., 2023) is a CLIP-based preference model trained on the Human Preference Dataset v2 (HPD v2), which contains large-scale pairwise human preference annotations. HPS v2 is designed to provide robust preference prediction across diverse image distributions and serves as a complementary measure of human-aligned generation quality.

AES. We use the Aesthetic Score (AES) (Schuhmann, 2022) to evaluate the perceptual attractiveness of generated images. AES is predicted from CLIP image representations using an aesthetic predictor trained on human-rated images, providing a prompt-independent measure of visual aesthetics. Unlike preference metrics that jointly consider text and image, AES primarily reflects the overall aesthetic quality of the generated image.

VisualQuality-R1. We further employ VisualQuality-R1 (Wu et al., 2026c), a reasoning-enhanced no-reference image quality assessment model trained with reinforcement learning to rank. VisualQuality-R1 evaluates perceptual image quality without requiring a reference image and provides a complementary assessment of visual fidelity beyond text–image preference metrics.

A.3 Model Backbones

We evaluate DA-DMD on five text-to-image diffusion backbones from two representative model families: SANA, SANA-1.5, Stable Diffusion 3 Medium (SD3-M), Stable Diffusion 3.5 Medium (SD3.5-M), and Stable Diffusion 3.5 Large (SD3.5-L). These models cover different architectures and model scales, allowing us to evaluate whether the proposed disagreement-aware optimization generalizes beyond a particular backbone.

SANA and SANA-1.5. SANA (Xie et al., 2024) is an efficient text-to-image framework based on a linear Diffusion Transformer architecture, designed to achieve high-resolution image generation with substantially reduced computational cost. SANA-1.5 (Xie et al., 2025) further scales the SANA family through improved training-time and inference-time scaling strategies. We include both generations of SANA to evaluate DA-DMD across closely related but differently scaled diffusion backbones.

Stable Diffusion 3 and 3.5. Stable Diffusion 3 (Esser et al., 2024) introduces the Multimodal Diffusion Transformer (MMDiT) architecture for text-to-image generation. We use SD3-M together with the subsequent SD3.5-M and SD3.5-L variants, covering both medium- and large-scale models. Evaluating DA-DMD across these three backbones allows us to study its behavior under different model capacities within the same general architecture family.

A.4 Hyperparameter Settings

Unless otherwise specified, DA-DMD follows the same training recipe as the corresponding DMD2 baseline. Following the multi-step distribution-matching and fake-model optimization framework of DMD2, we use only the score-based distribution-matching objective in all experiments and disable the real-data adversarial objective for both DMD2 and DA-DMD. Therefore, the only difference between the two methods is the disagreement-aware reweighting applied to the DMD update.

All models are trained at 1024×1024 resolution with a 4-step student. The teacher is frozen, while both the student and fake model are initialized from the corresponding pretrained backbone. We use AdamW with a learning rate of 5×10^{-6} for both the student and fake model, and perform five fake-model optimization steps per student update. For DA-DMD, the active high-noise region corresponds to the first 30% of the noise trajectory, i.e., normalized noise level $\sigma_t^{\text{norm}} \geq 0.7$; outside this region, the original DMD update is left unchanged. The main model-specific settings are summarized in Tab. 2.

For all benchmarks, distilled models are evaluated using 4 sampling steps at 1024×1024 resolution with guidance scale 1.0. We use the same evaluation seed and metric implementation for DMD2 and DA-DMD. The weighting coefficient α and intervention region are further analyzed in Sec. 3.3.

Table 2: **Main training hyperparameters.** DA-DMD uses the same optimization and sampling configuration as DMD2, with model-specific teacher guidance and disagreement weighting coefficient α.

Backbone	Teacher NFE	Teacher CFG	α
SANA	20	4.5	0.2
SANA-1.5	20	4.5	0.2
SD3-M	30	7.0	0.3
SD3.5-M	30	3.5	0.3
SD3.5-L	30	3.5	0.7

The coefficient α controls the maximum retained DMD update strength within the active high-noise region through $\omega_\alpha(0) = \alpha$. We do not assume a universal value across model families. Instead, α is selected independently for each backbone from a small candidate set using the same tuning protocol. Our ablations show that performance is sensitive to the amount of attenuation and that moderate retention generally performs well, while the preferred value can shift across architectures. We therefore treat α as a model-dependent optimization hyperparameter rather than attributing its variation to model scale or parameter count.

B Related Work

Few-step diffusion distillation. Diffusion and flow-based generative models achieve strong synthesis quality but typically require iterative sampling, motivating extensive work on reducing the number of function evaluations at inference time. Early approaches compress multi-step teachers through progressive distillation (Salimans & Ho, 2022), while consistency-based methods learn mappings that directly connect noisy states to clean samples or enforce consistency along probability-flow trajectories (Song et al., 2023; Luo et al., 2023). More recent methods explore trajectory-, transition-, and output-level distillation objectives to further reduce sampling steps while preserving generation quality (Luo et al., 2026; Zhou et al., 2024). These approaches primarily constrain the student through teacher-generated targets or trajectories. In contrast, distribution matching methods optimize the student at the level of its induced marginal distribution, allowing the student trajectory to deviate from the teacher's original sampling path.

Distribution matching distillation. Distribution Matching Distillation (DMD) (Yin et al., 2024b) introduced a distribution-level approach to few-step generation by minimizing an approximate KL divergence between teacher and student distributions. Its gradient is estimated through the difference between a pretrained teacher score and a "fake" score model trained to track the evolving student distribution. DMD2 (Yin et al., 2024a) substantially improves this framework by removing the original regression-data requirement, stabilizing fake-model estimation with a two-time-scale update rule, incorporating real-data adversarial supervision, and extending DMD to multi-step generation. Subsequent work has improved different aspects of the DMD framework. Diversity-Preserved DMD (Wu et al., 2026a) separates the roles of distilled steps to mitigate diversity collapse, while Decoupled DMD (Liu et al., 2025) decomposes the effects of classifier-free guidance augmentation and distribution matching and uses this decomposition to redesign the distillation schedule. Other studies investigate broader design choices surrounding distribution matching, including training recipes, teacher guidance, and data composition (Wu et al., 2026b). Together, these works demonstrate that both the objective and the surrounding optimization procedure are critical to successful few-step distribution matching.

Adaptive optimization of DMD. A growing line of work further studies when standard DMD supervision may become ineffective or unstable. Adaptive Matching Distillation (AMD) (Bai et al., 2026), for example, identifies failure regions in the DMD optimization landscape and employs reward-based proxies together with adaptive correction mechanisms to steer the student away from such regions. Other approaches introduce auxiliary perceptual, adversarial, reward, or task-specific signals to improve fidelity, diversity, or training stability (Jiang et al., 2025; Wu et al., 2026a). These

methods highlight the importance of adapting DMD optimization, but typically do so by augmenting the supervision, modifying individual objective components, or introducing additional mechanisms.

Our work is complementary to these directions. Rather than adding a new source of supervision or explicitly detecting failure states with an external proxy, we ask whether DMD's native teacher–fake discrepancy can itself diagnose the aggressiveness of its updates. We find that disagreement magnitude contains such information and use it as a continuous control signal for selective high-noise attenuation. DA-DMD therefore modifies how strongly existing DMD supervision is applied, rather than introducing a new target for the student.

C Additional Theoretical Analysis

This section provides the complete derivations and complementary interpretations underlying the theoretical results in Secs. 2.1–2.3. We first derive the score–clean prediction identity used to express the DMD update in terms of teacher–fake clean predictions. We then analyze disagreement-aware attenuation from optimization, robustness, and gradient-estimation perspectives, followed by proximal and probabilistic interpretations of the practical weighting rule.

C.1 Score–Clean Prediction Identity

We first derive Eq. 4. Consider the Gaussian perturbation process

$$x_t = \alpha_t x_0 + \sigma_t \epsilon, \qquad \epsilon \sim \mathcal{N}(0, I), \tag{17}$$

which induces the conditional transition

$$q_t(x_t \mid x_0) = \mathcal{N}\left(x_t; \alpha_t x_0, \sigma_t^2 I\right). \tag{18}$$

For a conditional clean-data distribution $p_0(x_0 \mid c)$, the noisy marginal is

$$p_t(x_t \mid c) = \int p_0(x_0 \mid c) q_t(x_t \mid x_0)\, dx_0. \tag{19}$$

Applying Fisher's identity gives

$$\begin{aligned} s_t(x_t, c) &= \nabla_{x_t} \log p_t(x_t \mid c) \\ &= \mathbb{E}\left[\nabla_{x_t} \log q_t(x_t \mid x_0) \mid x_t, c\right]. \end{aligned} \tag{20}$$

For the Gaussian transition in Eq. 18,

$$\nabla_{x_t} \log q_t(x_t \mid x_0) = -\frac{x_t - \alpha_t x_0}{\sigma_t^2}. \tag{21}$$

Substituting Eq. 21 into Eq. 20 yields

$$s_t(x_t, c) = \frac{\alpha_t \mathbb{E}[x_0 \mid x_t, c] - x_t}{\sigma_t^2}, \tag{22}$$

which recovers Eq. 4.

For an x_0-prediction model trained with squared-error denoising, the population-optimal predictor is

$$\hat{x}_0^\star(x_t, t, c) = \mathbb{E}[x_0 \mid x_t, c]. \tag{23}$$

Hence, a trained model satisfies $\hat{x}_0(x_t, t, c) \approx \mathbb{E}[x_0 \mid x_t, c]$, and Eq. 22 gives

$$s_t(x_t, t, c) \approx \frac{\alpha_t \hat{x}_0(x_t, t, c) - x_t}{\sigma_t^2}. \tag{24}$$

Since the teacher and fake model are evaluated at the same noisy sample x_t, the shared term $-x_t/\sigma_t^2$ cancels:

$$s_F - s_T \approx \frac{\alpha_t}{\sigma_t^2}\left(\hat{x}_{0,F} - \hat{x}_{0,T}\right). \tag{25}$$

Thus, the clean-prediction disagreement used by DA-DMD is directly proportional, up to timestep-dependent scaling, to the native teacher–fake score discrepancy driving the DMD correction.

C.2 One-Step Adaptive Attenuation

We now provide the complete derivation of the adaptive step-control result summarized in Sec. 2.3. Let $g = g_{\mathrm{DMD}}$, and consider a scaled update

$$\theta^+ = \theta - \eta a g, \qquad a \in [0, 1]. \tag{26}$$

Assume that the DMD objective $\mathcal{L}(\theta)$ is β-smooth. Then

$$\mathcal{L}(\theta^+) \le \mathcal{L}(\theta) - \eta a \langle \nabla \mathcal{L}(\theta), g \rangle + \frac{\beta \eta^2 a^2}{2} \|g\|^2. \tag{27}$$

Conditioned on noise level t and disagreement r, define

$$\begin{aligned} A(t, r) &= \mathbb{E}\left[\langle \nabla \mathcal{L}(\theta), g \rangle \mid t, r\right] > 0, \\ B(t, r) &= \mathbb{E}\left[\|g\|^2 \mid t, r\right] > 0. \end{aligned} \tag{28}$$

Here, $A(t, r)$ measures useful descent alignment, while $B(t, r)$ measures update energy. Taking the conditional expectation of Eq. 27 gives

$$\mathbb{E}\left[\mathcal{L}(\theta^+) \mid t, r\right] \le \mathcal{L}(\theta) - \eta a A(t, r) + \frac{\beta \eta^2 a^2}{2} B(t, r). \tag{29}$$

Optimal retained update strength. The a-dependent part of the bound is $Q(a) = -\eta a A + \frac{\beta \eta^2 a^2}{2} B$. Differentiating gives $\partial Q / \partial a = -\eta A + \beta \eta^2 a B$, whose unconstrained minimizer is $A/(\beta \eta B)$. Enforcing $a \in [0, 1]$ therefore yields

$$a^\star(t, r) = \min\left\{1, \frac{A(t, r)}{\beta \eta B(t, r)}\right\}. \tag{30}$$

Equivalently,

$$\frac{B(t, r)}{A(t, r)} \uparrow \quad \Longrightarrow \quad a^\star(t, r) \downarrow. \tag{31}$$

Eq. 31 formalizes the distinction between update direction and update magnitude: a DMD update may remain positively aligned with the descent direction while still being too aggressive to apply at full strength.

When is attenuation preferable to vanilla DMD? For a fixed attenuation $a \in (0, 1)$, let U_a denote the right-hand side of Eq. 29, and let U_1 denote the vanilla-DMD bound. Their difference is

$$\begin{aligned} U_a - U_1 &= \eta(1 - a) A - \frac{\beta \eta^2}{2}(1 - a^2) B \\ &= \eta(1 - a)\left[A - \frac{\beta \eta}{2}(1 + a) B\right]. \end{aligned} \tag{32}$$

Since $a \in (0, 1)$, we obtain

$$U_a < U_1 \quad \Longleftrightarrow \quad \frac{B(t, r)}{A(t, r)} > \frac{2}{\beta \eta (1 + a)}. \tag{33}$$

Therefore, attenuation admits a strictly tighter one-step smoothness bound whenever update energy is sufficiently large relative to useful descent alignment.

Connection to disagreement. Suppose that within the active high-noise region $\mathcal{H}$, the conditional energy-to-alignment ratio

$$\rho(t, r) = \frac{B(t, r)}{A(t, r)} \tag{34}$$

is nondecreasing with disagreement r. Combining this assumption with Eq. 30 gives

$$r \uparrow \quad \Longrightarrow \quad a^\star(t, r) \downarrow. \tag{35}$$

This monotonic trend is consistent with the practical DA-DMD weighting rule in Eq. 13, where larger teacher–fake disagreement induces stronger attenuation. Importantly, this argument does not require high-disagreement updates to be incorrect; it only requires their update energy to grow sufficiently relative to useful descent alignment.

C.3 Selective Energy Control and Redescending Influence

We next establish the selective nature of DA-DMD and characterize its response to extreme disagreement.

Selective update-energy control. Ignoring the stop-gradient operator, which does not change the forward value of the multiplier, Eq. 14 gives

$$g_{\mathrm{DA-DMD}} = [1 - \chi + \chi\omega_\alpha(r)]\, g_{\mathrm{DMD}}, \qquad \chi = \mathbf{1}[t \in \mathcal{H}]. \tag{36}$$

Since $\chi \in \{0, 1\}$,

$$[1 - \chi + \chi\omega_\alpha(r)]^2 = 1 - \chi\left[1 - \omega_\alpha^2(r)\right]. \tag{37}$$

Combining Eqs. 36 and 37 yields

$$\begin{aligned}\mathbb{E}\|g_{\mathrm{DA-DMD}}\|^2 = \mathbb{E}\|g_{\mathrm{DMD}}\|^2 \\ - \mathbb{E}\left[\mathbf{1}[t \in \mathcal{H}]\left(1 - \omega_\alpha^2(r)\right)\|g_{\mathrm{DMD}}\|^2\right].\end{aligned} \tag{38}$$

Hence, the instantaneous update-energy reduction induced by DA-DMD is localized entirely to the active region. In particular,

$$t \notin \mathcal{H} \quad \Longrightarrow \quad g_{\mathrm{DA-DMD}} = g_{\mathrm{DMD}}. \tag{39}$$

Monotonic and redescending influence. Recall the practical weighting rule

$$\omega_\alpha(r) = \frac{\alpha e^{-r^2/2}}{\alpha e^{-r^2/2} + 1 - \alpha}. \tag{40}$$

Its derivative is

$$\frac{d\omega_\alpha}{dr} = -r\,\omega_\alpha(r)\left[1 - \omega_\alpha(r)\right], \tag{41}$$

which is strictly negative for $r > 0$. Thus, larger disagreement always induces stronger attenuation.

Moreover,

$$\omega_\alpha(r) \le \frac{\alpha}{1 - \alpha} e^{-r^2/2}. \tag{42}$$

Multiplying both sides by $r \ge 0$ gives

$$0 \le r\omega_\alpha(r) \le \frac{\alpha}{1 - \alpha} r e^{-r^2/2}. \tag{43}$$

Since $re^{-r^2/2} \to 0$ as $r \to \infty$,

$$\boxed{\lim_{r\to\infty} r\,\omega_\alpha(r) = 0.} \tag{44}$$

The disagreement-dependent correction therefore exhibits redescending influence: increasingly extreme teacher–fake disagreement receives progressively diminishing effective weight instead of unbounded influence.

Ideal fixed-point preservation. At the ideal distribution-matching solution, the student matches the teacher and the fake model perfectly tracks the student. Hence $s_F = s_T$, which implies $g_{\mathrm{DMD}} = 0$. Eq. 14 therefore gives

$$g_{\mathrm{DMD}} = 0 \quad \Longrightarrow \quad g_{\mathrm{DA-DMD}} = 0. \tag{45}$$

Thus, DA-DMD modifies the optimization dynamics while preserving the ideal distribution-matching solution. We do not require or claim that all non-ideal stationary points of DA-DMD and vanilla DMD are identical.

C.4 Bias–Variance Interpretation

The smoothness analysis above studies the quality of a single optimization step. A complementary view interprets disagreement-aware attenuation as conditional shrinkage of a stochastic DMD update.

Suppose

$$g = g^\star + \epsilon, \qquad \mathbb{E}[\epsilon \mid t, r] = 0, \tag{46}$$

where $g^\star$ denotes the conditional mean update. Define the conditional stochastic variation as

$$V(t, r) = \mathbb{E}\left[\|\epsilon\|^2 \mid t, r\right]. \tag{47}$$

The vanilla estimator therefore has $\mathrm{MSE}_{\mathrm{DMD}} = V$.

For a shrunk estimator $g_a = ag$, $a \in (0, 1)$, $g_a - g^\star = (a-1)g^\star + a\epsilon$. Using the conditional zero-mean assumption gives

$$\mathrm{MSE}(a) = (1-a)^2\|g^\star\|^2 + a^2 V. \tag{48}$$

The first term is the bias introduced by shrinkage, while the second is the remaining stochastic variation.

Comparing Eq. 48 with the vanilla estimator yields

$$\mathrm{MSE}(a) < \mathrm{MSE}_{\mathrm{DMD}} \quad \Longleftrightarrow \quad V > \frac{1-a}{1+a}\|g^\star\|^2. \tag{49}$$

Minimizing Eq. 48 with respect to a gives

$$a^\star_{\mathrm{MSE}} = \frac{\|g^\star\|^2}{\|g^\star\|^2 + V}. \tag{50}$$

Therefore,

$$\frac{V}{\|g^\star\|^2} \uparrow \quad \Longrightarrow \quad a^\star_{\mathrm{MSE}} \downarrow. \tag{51}$$

This bias–variance view complements the smoothness analysis in Sec. C.2: both favor more conservative updates when the stochastic DMD correction becomes large or uncertain relative to its useful signal.

We emphasize that this interpretation does not assume that disagreement r directly estimates $V(t, r)$; rather, it provides an independent statistical explanation for why conditional shrinkage can improve a noisy update estimator.

C.5 Proximal and Energy-Based Interpretations

We finally provide two complementary interpretations of the practical disagreement-aware weighting rule.

Proximal interpretation. Within the active region $t \in \mathcal{H}$, DA-DMD scales the original teacher-directed correction by $\omega_\alpha(r)$. Define the adaptive proximal target

$$\hat{x}_{0,P} = [1 - \omega_\alpha(r)]\,\hat{x}_{0,F} + \omega_\alpha(r)\hat{x}_{0,T}. \tag{52}$$

Then

$$\hat{x}_{0,F} - \hat{x}_{0,P} = \omega_\alpha(r)\,(\hat{x}_{0,F} - \hat{x}_{0,T}). \tag{53}$$

Thus, scaling the DMD correction by $\omega_\alpha(r)$ is equivalent to replacing the full teacher-directed correction with a shorter correction toward an adaptive point between the fake and teacher predictions. As disagreement increases, $\omega_\alpha(r)$ decreases and the effective target moves toward $\hat{x}_{0,F}$, yielding a more conservative update. This is the geometric interpretation illustrated in Fig. 3.

Energy-based responsibility interpretation. The same weighting rule also admits an interpretation in terms of normalized energy-based responsibilities. Consider two disagreement-dependent compatibility potentials associated with the teacher-directed and fake-centered components:

$$\psi_T(r) = e^{-r^2/2}, \qquad \psi_F(r) = 1. \tag{54}$$

Together with mixture coefficients $\pi_T = \alpha$ and $\pi_F = 1-\alpha$, we define the normalized responsibility of the teacher-directed component as

$$\rho_T(r) = \frac{\pi_T \psi_T(r)}{\pi_T \psi_T(r) + \pi_F \psi_F(r)} = \frac{\alpha e^{-r^2/2}}{\alpha e^{-r^2/2} + 1 - \alpha} = \omega_\alpha(r). \tag{55}$$

Accordingly, $\rho_F(r) = 1 - \rho_T(r)$. As disagreement increases, the compatibility of the teacher-directed component decreases relative to the fake-centered component, resulting in a progressively smaller teacher responsibility.

Using these responsibilities to interpolate between the two predictions gives

$$\rho_T(r)\hat{x}_{0,T} + \rho_F(r)\hat{x}_{0,F} = \omega_\alpha(r)\hat{x}_{0,T} + [1 - \omega_\alpha(r)]\,\hat{x}_{0,F}, \tag{56}$$

which exactly recovers the proximal target in Eq. 52.

Importantly, ψ_T and ψ_F are unnormalized compatibility potentials rather than probability densities over r. Therefore, Eq. 55 should be interpreted as an energy-based normalization of relative responsibilities, rather than as an estimate of the true teacher–fake density ratio.

C.6 Arithmetic versus Geometric Mixtures

We briefly compare two natural interpolations between the teacher and fake distributions and explain why DA-DMD adopts the arithmetic mixture.

Arithmetic mixture. Consider

$$p^{\text{arith}}_{\alpha,t}(x_t) = \alpha p_{T,t}(x_t) + (1-\alpha) p_{F,t}(x_t). \tag{57}$$

Its score is

$$s^{\text{arith}}_\alpha = \rho_\alpha(x_t) s_T + [1 - \rho_\alpha(x_t)]\, s_F, \tag{58}$$

where

$$\rho_\alpha(x_t) = \frac{\alpha p_{T,t}(x_t)}{\alpha p_{T,t}(x_t) + (1-\alpha) p_{F,t}(x_t)} = \mathcal{W}_\alpha(\Delta_t(x_t)). \tag{59}$$

Therefore,

$$s_F - s^{\text{arith}}_\alpha = \rho_\alpha(x_t)\,(s_F - s_T)\,. \tag{60}$$

Although the mixture coefficient α is fixed globally, the effective score-space coefficient $\rho_\alpha(x_t)$ depends on the sample. The arithmetic mixture therefore naturally induces sample-adaptive modulation of the original DMD correction.

Geometric mixture. Alternatively, consider the normalized geometric interpolation

$$p^{\text{geo}}_{\alpha,t}(x_t) = \frac{p_{T,t}(x_t)^\alpha p_{F,t}(x_t)^{1-\alpha}}{Z_\alpha}, \tag{61}$$

where Z_α is the normalization constant. Since Z_α does not depend on x_t,

$$\begin{aligned} s^{\text{geo}}_\alpha &= \nabla_{x_t} \log p^{\text{geo}}_{\alpha,t}(x_t) \\ &= \alpha s_T + (1-\alpha) s_F. \end{aligned} \tag{62}$$

Hence,

$$s_F - s^{\text{geo}}_\alpha = \alpha\,(s_F - s_T)\,. \tag{63}$$

The geometric mixture therefore preserves the DMD correction direction but only rescales it by the constant factor α, independent of the sample.

Implication for DA-DMD. Both constructions produce intermediate distributions between the teacher and fake models, but they differ critically in score space. The geometric mixture yields uniform attenuation, whereas the arithmetic mixture converts a fixed distribution-level mixture coefficient into a sample-dependent score-space responsibility. Since our goal is to adapt the update strength according to sample-level training signals while preserving the original DMD correction direction, the arithmetic mixture provides the more appropriate structural template.

We emphasize that this comparison motivates the *form* of sample-adaptive weighting. The practical DA-DMD weight further replaces the unavailable density-ratio-dependent responsibility with the disagreement-based control rule introduced in Sec. 2.2.

D Additional Experiments

We provide additional experiments complementing the main-paper evaluation, including controlled analyses of DMD optimization, complete quantitative results on additional benchmarks, and qualitative comparisons across all evaluated backbones.

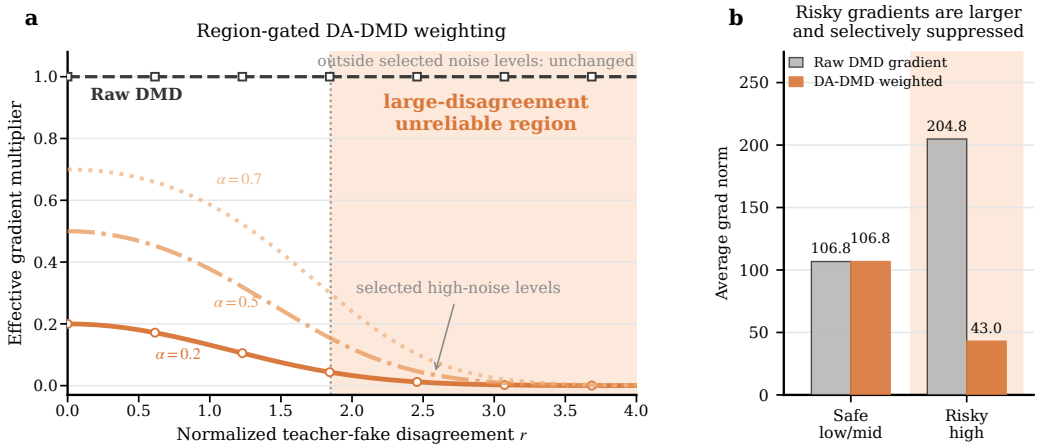

Figure 7: **Selective disagreement-aware weighting in DA-DMD.** (a) Outside the predefined high-noise region $\mathcal{H}$, DA-DMD exactly preserves the vanilla DMD update. Within $\mathcal{H}$, the retained update strength decreases continuously with normalized teacher–fake disagreement r, with α controlling its maximum value. (b) The resulting weighting leaves low-/mid-noise updates unchanged while substantially attenuating the high-energy updates concentrated in the high-noise regime.

D.1 Diagnostic Analysis

We first provide controlled analyses to further understand the relationship between teacher–fake disagreement and DMD optimization. Figs. 8, 9, and 10 show results on DIGITS, CIFAR-10, and CIFAR-100, respectively. Starting from the same initialization, high-disagreement samples consistently exhibit larger raw DMD updates and can disproportionately influence the optimization trajectory, leading to biased target-mode occupancy. DA-DMD attenuates these aggressive corrections and recovers the target distribution more faithfully across all three settings.

Fig. 7 further illustrates the selective weighting mechanism of DA-DMD. Updates outside the predefined high-noise region remain exactly unchanged, whereas disagreement continuously controls the retained update strength within the active region. Together with the large-scale SD3.5-M analysis in Fig. 2, these results consistently support the central design principle of DA-DMD: noise level determines *where* intervention is useful, while teacher–fake disagreement determines *how strongly* the update should be moderated.

D.2 Additional Quantitative Results

PartiPrompts. Tab. 3 reports the complete results on PartiPrompts. DA-DMD achieves higher average pairwise win rates than DMD2 across all five backbones, demonstrating that its improvements extend beyond DrawBench to a substantially larger and more diverse prompt set. The gains are particularly pronounced on SD3.5-M and SD3.5-L, where DA-DMD reaches average win rates of 61.2% and 61.4%, respectively.

GenEval2. Tabs. 4 and 5 evaluate compositional generation from complementary perspectives. DA-DMD improves the overall GenEval2 score across the evaluated backbones, with the largest gain of $+6.81$ on SD3.5-L. Improvements become especially pronounced on more compositionally demanding prompts: on SD3.5-L, DA-DMD improves the 7- and 8-atom scores by $+13.04$ and $+9.56$, respectively. These results suggest that disagreement-aware update control is particularly beneficial when multiple visual constraints must be satisfied simultaneously.

TIIF-Bench. Tabs. 6 and 7 provide a detailed evaluation of fine-grained instruction following. DA-DMD improves the overall TIIF-Bench score across all five evaluated backbones, with the largest gain of $+2.44$ on SD3.5-L. Category-level performance varies across individual skills, but the aggregate results show broad improvements on both short- and long-form instructions, including advanced compositional settings.

D.3 Additional Qualitative Results

We further compare DA-DMD with DMD2 across all evaluated model families. Figs. 12, 13, 14, 15, and 16 present additional examples on SD3-M, SD3.5-M, SD3.5-L, SANA, and SANA-1.5, respectively. Across diverse prompts involving counting, spatial relations, relative attributes, color binding, text rendering, and structured multi-object composition, DA-DMD more faithfully satisfies the specified visual constraints while maintaining comparable perceptual quality. These examples complement the quantitative results and show that the improvements produced by disagreement-aware update control are also reflected in visually observable prompt adherence.

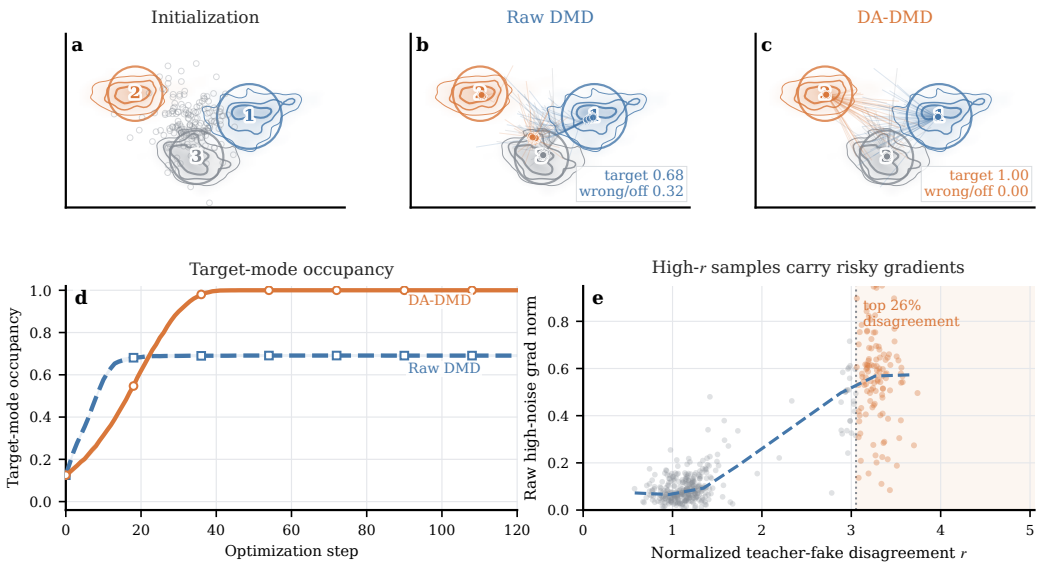

Figure 8: **Controlled analysis of disagreement-aware DMD optimization on CIFAR-10.** Starting from the same initialization, raw DMD can be dominated by high-disagreement samples with large update magnitudes, leading to an imbalanced optimization trajectory. DA-DMD selectively attenuates these updates while retaining lower-disagreement corrections, resulting in improved target-mode matching. The bottom panels visualize the evolution of target-mode occupancy and the concentration of large raw DMD updates among high-disagreement samples.

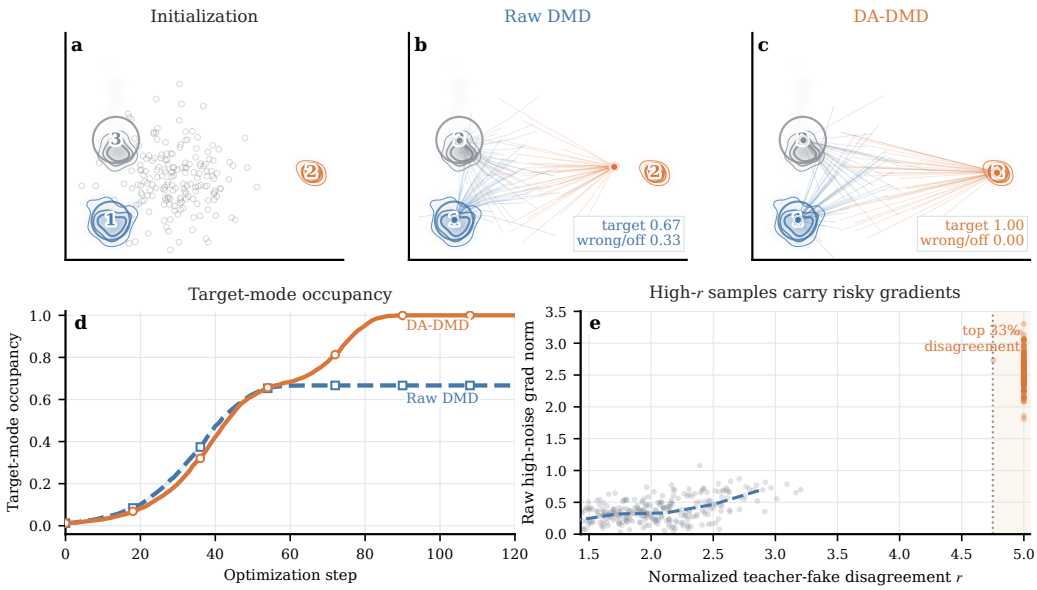

Figure 9: **Controlled analysis of disagreement-aware DMD optimization on DIGITS.** Starting from the same initialization, raw DMD can be dominated by high-disagreement samples with large update magnitudes, leading to an imbalanced optimization trajectory. DA-DMD selectively attenuates these updates while retaining lower-disagreement corrections, resulting in improved target-mode matching. The bottom panels visualize the evolution of target-mode occupancy and the concentration of large raw DMD updates among high-disagreement samples.

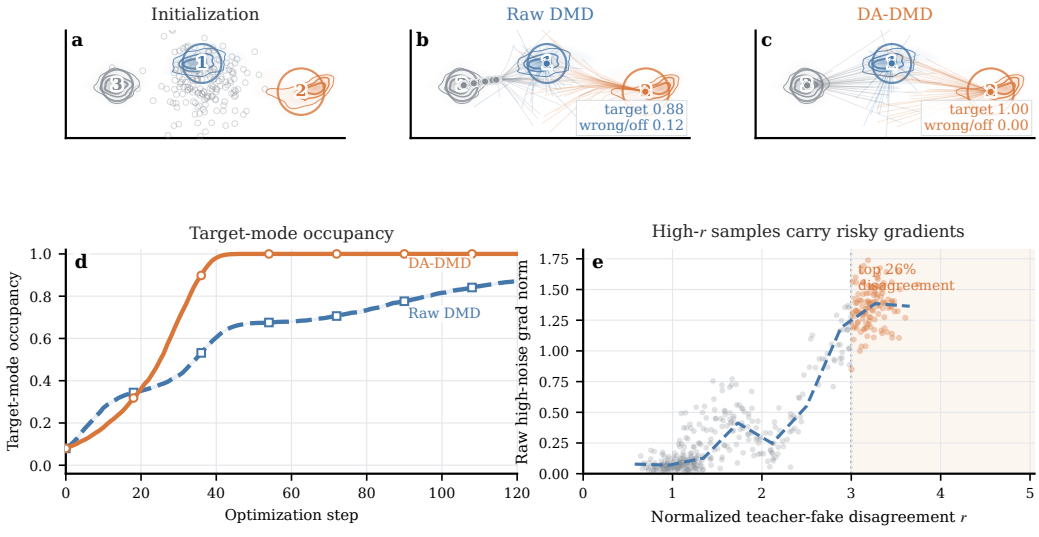

Figure 10: **Controlled analysis of disagreement-aware DMD optimization on CIFAR-100.** Starting from the same initialization, raw DMD can be dominated by high-disagreement samples with large update magnitudes, leading to an imbalanced optimization trajectory. DA-DMD selectively attenuates these updates while retaining lower-disagreement corrections, resulting in improved target-mode matching. The bottom panels visualize the evolution of target-mode occupancy and the concentration of large raw DMD updates among high-disagreement samples.

Table 3: **Quantitative comparison on PartiPrompts.** DA-DMD yields broad improvements over the corresponding DMD2 baselines across different model families.

Backbone	Method	PickScore↑	ImageReward↑	HPSv2↑	AES↑	VQ-R1↑	Avg. WR↑
SANA	DMD2	22.74	117.28	**30.57**	5.940	4.617	-
	DA-DMD (Ours)	**22.77** +0.03	**119.58** +2.30	**30.57** +0.00	**5.956** +0.016	**4.638** +0.020	**53.6%**
SANA-1.5	DMD2	22.85	**122.79**	30.31	6.022	4.553	-
	DA-DMD (Ours)	**22.90** +0.05	122.29 -0.50	**30.44** +0.13	**6.050** +0.029	**4.575** +0.023	**54.6%**
SD3-M	DMD2	22.39	116.74	30.44	5.697	4.571	-
	Hyper-SD3	22.24	92.07	27.87	5.530	4.418	30.0%
	SiD-DiT	22.56	109.77	28.98	5.499	4.617	38.8%
	SiD-Adv	22.53	104.79	29.86	5.716	4.583	44.4%
	TDM	22.54	**117.95**	30.38	5.647	4.648	50.2%
	DA-DMD (Ours)	**22.57** +0.17	113.96 -2.78	**30.84** +0.40	**5.726** +0.030	**4.678** +0.107	**57.2%**
SD3.5-M	DMD2	22.69	119.55	29.56	5.543	4.611	-
	SiD-DiT	22.71	114.97	29.87	5.661	4.661	49.0%
	SiD-Adv	22.67	114.72	29.45	**5.816**	4.553	48.4%
	DP-DMD	22.48	114.12	28.83	5.486	4.560	36.6%
	SenseFlow	22.50	104.05	30.18	5.714	4.674	56.2%
	DA-DMD (Ours)	**22.75** +0.06	**123.73** +4.18	**30.33** +0.77	5.659 +0.116	**4.692** +0.081	**61.2%**
SD3.5-L	DMD2	22.69	116.32	29.14	5.595	4.591	-
	SiD-DiT	22.81	112.08	29.95	5.683	4.621	54.0%
	SiD-Adv	22.98	118.46	30.01	**5.849**	4.628	61.2%
	SD3.5-L Turbo	22.85	110.84	29.14	5.755	4.616	52.4%
	SenseFlow	22.89	119.52	30.48	5.639	**4.671**	60.8%
	DA-DMD (Ours)	**23.10** +0.41	**121.78** +5.46	**30.53** +1.39	5.694 +0.099	4.588 -0.003	**61.4%**

Table 4: **GenEval2 results across backbones.** Δ denotes the absolute improvement over DMD2. WR is the pairwise win rate of DA-DMD against DMD2.

Backbone	Overall ↑			WR (%) ↑
	DMD2	DA-DMD	Δ	
SANA	14.68	**15.89**	+1.21	55.0
SD3.5-M	20.17	**21.74**	+1.57	55.4
SD3.5-L	15.93	**22.74**	+6.81	61.9

Table 5: **GenEval2 results across prompt complexity.** DA-DMD provides broad gains across compositional complexity levels, with particularly strong improvements on challenging high-atom prompts, e.g., +13.04 and +9.56 on 7- and 8-atom prompts with SD3.5-L. Colored values denote absolute changes over DMD2.

Backbone	Method	3↑	4↑	5↑	6↑	7↑	8↑	9↑	10↑
SANA	DMD2	42.03	**32.25**	**14.39**	11.27	6.48	5.53	3.22	2.23
	DA-DMD	**44.68** +2.65	32.11 -0.14	13.89 -0.50	**13.08** +1.81	**10.41** +3.93	**7.08** +1.55	**4.22** +1.00	**3.37** +1.14
SANA-1.5	DMD2	53.69	43.07	20.28	**18.02**	**6.95**	**9.27**	5.63	3.43
	DA-DMD	**54.66** +0.97	**44.20** +1.13	**22.62** +2.34	17.97 -0.05	6.59 -0.36	8.84 -0.43	**6.73** +1.10	**3.98** +0.55
SD3.5-M	DMD2	50.01	**39.16**	26.65	18.27	8.44	7.75	7.90	3.14
	DA-DMD	**53.47** +3.46	36.63 -2.53	**27.68** +1.03	**19.05** +0.78	**15.15** +6.71	**9.02** +1.27	**8.58** +0.68	**4.32** +1.18
SD3.5-L	DMD2	39.87	31.37	20.27	14.98	6.92	5.83	5.16	3.02
	DA-DMD	**45.83** +5.96	**42.11** +10.74	**26.81** +6.54	**20.10** +5.12	**19.96** +13.04	**15.39** +9.56	**7.81** +2.65	**3.89** +0.87

Table 6: **TIIF-Bench overall and basic-category results.** DA-DMD improves the overall instruction-following score across all five evaluated backbones, with the largest gain of +2.44 on SD3.5-L, while showing broadly positive improvements across short- and long-form categories.

Backbone	Method	Overall↑	Overall S↑	Overall L↑	Basic S↑	Basic L↑	Attr. S↑	Attr. L↑	Rela. S↑	Rela. L↑	Reason S↑	Reason L↑
SANA	DMD2	60.72	60.1	61.3	**75.8**	**77.3**	**79.3**	79.6	84.1	**84.6**	**63.4**	**67.6**
	DA-DMD	**61.12** +0.40	**60.4** +0.3	**61.8** +0.5	75.7 -0.1	77.1 -0.2	79.1 -0.2	**79.8** +0.2	**84.2** +0.1	84.3 -0.3	**63.4** +0.0	66.9 -0.7
SANA-1.5	DMD2	62.47	63.2	61.8	77.1	76.4	**81.2**	**79.2**	84.5	83.9	65.1	65.7
	DA-DMD	**62.84** +0.37	**63.5** +0.3	**62.2** +0.4	**77.5** +0.4	**76.6** +0.2	80.8 -0.4	79.1 -0.1	**84.8** +0.3	**84.7** +0.8	**66.3** +1.2	**66.1** +0.4
SD3-M	DMD2	63.33	63.8	62.8	77.8	77.8	83.5	82.5	**82.3**	**83.1**	66.4	**66.9**
	DA-DMD	**64.23** +0.90	**64.5** +0.7	**63.9** +1.1	**78.0** +0.2	**78.4** +0.6	**83.9** +0.4	**85.1** +2.6	82.0 -0.3	81.9 -1.2	**66.9** +0.5	66.7 -0.2
SD3.5-M	DMD2	66.77	**67.9**	65.6	**79.5**	79.3	**85.9**	84.0	84.4	**85.5**	**66.8**	67.6
	DA-DMD	**66.98** +0.21	67.5 -0.4	**66.5** +0.9	78.7 -0.8	**79.8** +0.5	84.3 -1.6	**85.2** +1.2	**85.2** +0.8	85.4 -0.1	65.7 -1.1	**68.0** +0.4
SD3.5-L	DMD2	65.82	68.4	63.3	77.8	75.6	84.6	80.2	83.3	81.4	64.2	64.3
	DA-DMD	**68.26** +2.44	**70.4** +2.0	**66.1** +2.8	**79.7** +1.9	**77.8** +2.2	**86.5** +1.9	**82.2** +2.0	**84.4** +1.1	**83.1** +1.7	**66.7** +2.5	**67.1** +2.8

Table 7: **TIIF-Bench advanced compositional results.** DA-DMD generally improves complex compositional instruction following across model backbones, with particularly strong gains on SD3.5-L, including +2.4 and +3.3 on advanced short- and long-form prompts. Colored values denote absolute changes over DMD2.

Backbone	Method	Adv. S↑	Adv. L↑	Attr.+Rela. S↑	Attr.+Rela. L↑	Attr.+Rea. S↑	Attr.+Rea. L↑	Rela.+Rea. S↑	Rela.+Rea. L↑
SANA	DMD2	51.8	52.8	78.0	76.3	63.3	**66.9**	66.0	68.7
	DA-DMD	**52.3** +0.5	**53.8** +1.0	**78.2** +0.2	**78.2** +1.9	**63.5** +0.2	**66.9** +0.0	**67.0** +1.0	**71.0** +2.3
SANA-1.5	DMD2	56.1	54.0	80.4	76.7	68.7	65.9	**68.8**	**69.4**
	DA-DMD	**56.3** +0.2	**54.7** +0.7	**80.6** +0.2	**76.8** +0.1	**69.8** +1.1	**67.2** +1.3	68.4 -0.4	68.6 -0.8
SD3-M	DMD2	57.1	55.4	78.0	78.7	63.9	65.7	63.2	66.9
	DA-DMD	**58.2** +1.1	**57.0** +1.6	**78.8** +0.8	**79.2** +0.5	**65.5** +1.6	**66.9** +1.2	**65.2** +2.0	**68.6** +1.7
SD3.5-M	DMD2	**62.5**	58.9	**80.9**	78.7	**68.2**	**67.4**	66.4	67.6
	DA-DMD	61.9 -0.6	**59.7** +0.8	79.9 -1.0	**78.8** +0.1	67.0 -1.2	67.1 -0.3	**66.8** +0.4	**68.8** +1.2
SD3.5-L	DMD2	64.0	57.0	80.1	75.6	65.1	64.4	63.5	62.8
	DA-DMD	**66.4** +2.4	**60.3** +3.3	**80.8** +0.7	**78.5** +2.9	**68.0** +2.9	**65.8** +1.4	**66.6** +3.1	**65.4** +2.6



Figure 11: **Qualitative comparison with representative distillation baselines on SD3.5-M.** DA-DMD demonstrates stronger prompt adherence and visual consistency across challenging cases involving counting, text rendering, fine-grained attributes, and object identity, compared with DMD, SiD-Adv, SiD-DiT, SenseFlow, and DP-DMD.

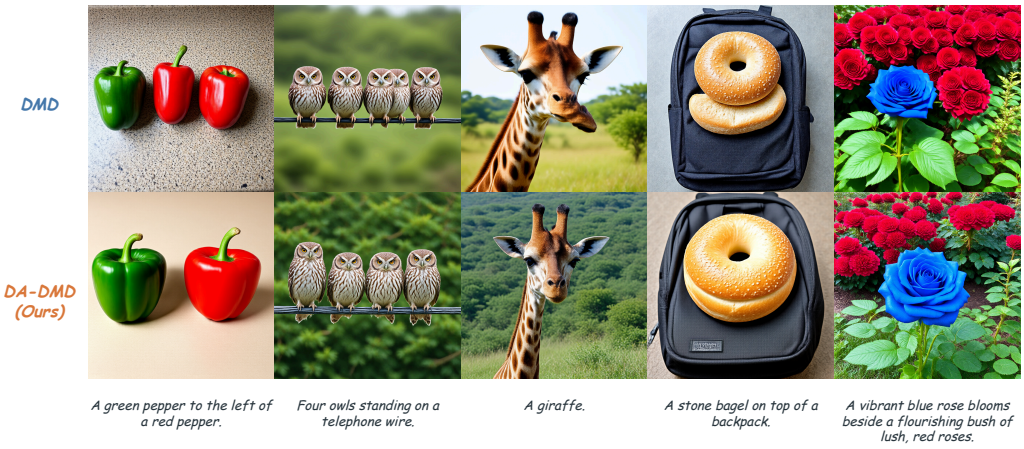

Figure 12: **Qualitative comparison on SD3-M.** DA-DMD improves prompt adherence over DMD2 across diverse compositional challenges, including object counting, spatial placement, relative relations, color control, and structured multi-object generation.



Figure 13: **Qualitative comparison on SD3.5-M.** DA-DMD improves prompt adherence over DMD2 across diverse compositional challenges, including object counting, spatial placement, relative relations, color control, and structured multi-object generation.



Figure 14: **Qualitative comparison on SD3.5-L.** DA-DMD improves prompt adherence over DMD2 across diverse compositional challenges, including object counting, spatial placement, relative relations, color control, and structured multi-object generation.

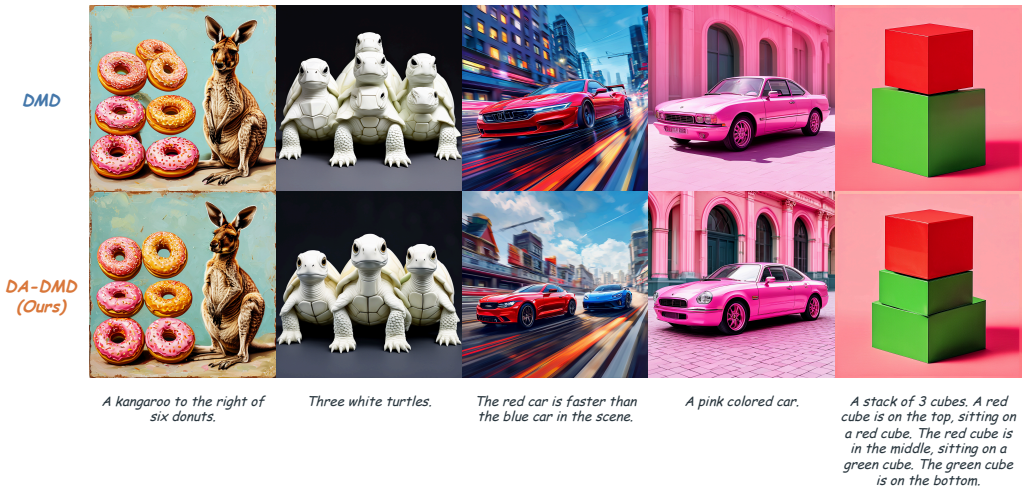

Figure 15: **Qualitative comparison on SANA.** DA-DMD improves prompt adherence over DMD2 across diverse compositional challenges, including object counting, spatial placement, relative relations, color control, and structured multi-object generation.



Figure 16: **Qualitative comparison on SANA-1.5.** DA-DMD improves prompt adherence over DMD2 across diverse compositional challenges, including object counting, spatial placement, relative relations, color control, and structured multi-object generation.