RESPONSE-DISAGREEMENT PRECONDITIONING FOR CAUSAL VIDEO DISTILLATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Distilling a bidirectional video diffusion model into a few-step causal generator requires transferring text-conditioned behavior across fundamentally different temporal factorizations. The teacher denoises with full-sequence context, whereas the student makes sequential, prefix-conditioned predictions. In asymmetric distribution matching distillation (DMD), the causal student is updated using the discrepancy between an online estimate of its rollout score and a classifier-free-guided teacher score. This discrepancy provides the base correction for distribution matching, but does not explicitly characterize that correction relative to differences in how the teacher and student score fields respond to text conditioning. We introduce Null-Relative Response-Preconditioned DMD (NRP-DMD) to exploit this response disagreement within the existing DMD update. At the same noised student rollout and diffusion timestep, we evaluate both score models under the target text and a null condition. For each model, the null-to-text change defines a text-induced score response, and the disagreement between the teacher and student responses defines a sample-dependent axis in score space. NRP-DMD uses this axis to precondition the existing guided DMD update, increasing the gain of its response-aligned component while leaving the orthogonal component unchanged. NRP-DMD modifies only the distillation procedure and requires neither architectural changes nor additional inference-time computation. Across two causal distillation baselines, NRP-DMD consistently improves generation quality over both 5- and 30-second horizons while maintaining strong text alignment and temporal coherence.

1 INTRODUCTION

State-of-the-art video diffusion models achieve strong visual fidelity and temporal coherence by processing the complete spatiotemporal sequence with bidirectional attention (Wan Team et al., 2025; Kong et al., 2024; Gao et al., 2025b). Such bidirectional computation requires access to the entire noisy sequence at each denoising step, making it ill-suited to streaming and interactive applications, where video content must be generated with low latency before future content is available. Causal video generators address this constraint by generating autoregressively and reusing cached representations of the past. Combined with few-step distillation, this formulation enables low-latency generation and has motivated a growing line of work that distills pretrained bidirectional video diffusion models into causal students (Yin et al., 2025; Huang et al., 2025; Zhu et al., 2026; Zhao et al., 2026).

Recent progress in bidirectional-to-causal transfer has strengthened the initialization, rollout, and supervision processes used to train a causal student. Asymmetric distribution matching distillation (DMD) provides a framework for learning from a bidirectional teacher (Yin et al., 2025). Within this framework, causal initialization determines the starting point for refinement, whereas self-rollout training aligns the histories encountered during optimization with those produced at inference (Zhu et al., 2026; Huang et al., 2025). Complementary developments extend these mechanisms to scalable few-step initialization (Zhao et al., 2026), real-time long-horizon streaming (Liu et al., 2026; Yang et al., 2026), and future-aware supervision of causal representations and historical context (Yu et al., 2026; Zhuang et al., 2026). These advances establish how the student's training trajectories and temporal context are constructed. Our focus is on the distribution-matching correction evaluated on

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.

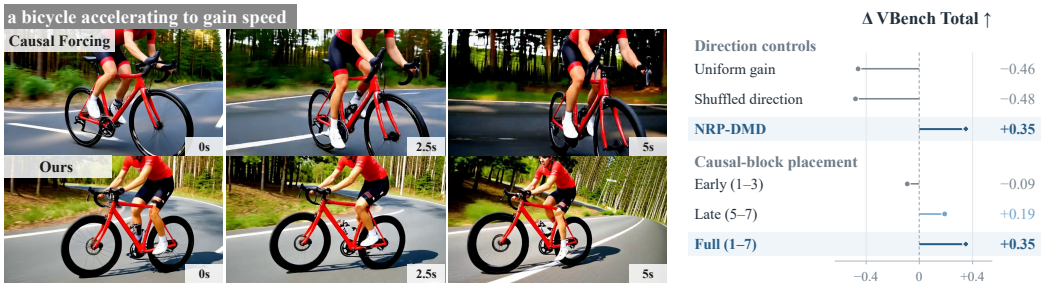

Figure 1: **NRP-DMD improves causal rollout quality through response-disagreement preconditioning.** Left: Frames sampled at 0, 2.5, and 5 seconds under the same prompt. Compared with Causal Forcing, NRP-DMD maintains more coherent subject geometry and scene structure throughout the causal rollout. Right: Changes in VBench Total relative to the DMD baseline. Neither uniformly scaling the DMD signal nor replacing the response-disagreement direction with a shuffled direction reproduces the gain of NRP-DMD. Applying the preconditioner to late causal blocks is more effective than restricting it to early blocks, while applying it to all blocks performs best.

these student rollouts: we study how to directionally reweight this correction while retaining the baseline causal rollout construction and the CFG-guided teacher target.

Specifically, we examine differences in how the teacher and estimated student-rollout score fields respond to the same text condition. In the asymmetric DMD update considered here, the guided score discrepancy already provides a sample-dependent correction between the estimated student-rollout distribution and the CFG-guided teacher target. At the same noised rollout and diffusion timestep, the change in each score field from a null reference to the target text defines its text-induced score response. The difference between the two responses defines a sample-dependent response-disagreement axis in score space. This disagreement measures how the unguided fake–teacher score residual changes from the null condition to the target text. We investigate whether this response-derived direction can guide a sample- and block-dependent reweighting of the existing DMD correction.

This response-derived axis is particularly relevant in bidirectional-to-causal distillation because the teacher and student realize text-conditioned behavior under different temporal information structures. The bidirectional teacher forms its denoising response while jointly processing the complete noisy video, whereas the causal student realizes the same text condition through a sequence of predictions conditioned on its generated prefix. Consequently, the student's conditional behavior depends jointly on the prompt and its evolving rollout state, while the teacher responds under full-sequence context. The same text condition can therefore induce different local score responses under the two temporal factorizations, motivating us to examine their disagreement at the rollout states where distribution matching is performed.

Building on this view, we introduce Null-Relative Response-Preconditioned DMD (NRP-DMD). NRP-DMD uses the response-disagreement axis to precondition the existing DMD update at each causal block, increasing the gain of the response-aligned component while leaving the orthogonal component unchanged. The resulting identity-plus-rank-one transformation preserves the pointwise zeros of the base DMD discrepancy.

Across two causal distillation frameworks, NRP-DMD improves generation quality from short clips to long-horizon rollouts while preserving prompt fidelity and temporal coherence. Controlled diagnostics in Figure 1 further show that neither matched uniform amplification nor a shuffled response direction reproduces these gains, supporting the relevance of the sample-dependent response direction. Because NRP-DMD acts only during distillation, it leaves the causal generator architecture and inference-time computation unchanged.

Our contributions are summarized as follows:

- We identify a sample-dependent response-disagreement direction in bidirectional-to-causal video distillation by comparing how the teacher and estimated student-rollout score fields change from a shared null condition to the target text at the same noised rollout and diffusion timestep.
- We introduce NRP-DMD, which uses the response-disagreement direction as an axis for preconditioning the existing guided DMD signal. The resulting blockwise identity-plus-rank-one preconditioner increases the gain of the response-aligned component while preserving the orthogonal component and the pointwise zeros of the base DMD discrepancy.
- We validate NRP-DMD on Self Forcing and Causal Forcing over both 5- and 30-second generation horizons, achieving higher VBench Total and Quality scores in all four settings without changing the causal generator architecture or adding inference-time computation.

2 Related Work

Causal video generation and distillation. Autoregressive video diffusion generates frames or latent chunks sequentially and reuses past representations as causal context (Chen et al., 2024; Gao et al., 2025a). CausVid distills a bidirectional teacher into a few-step causal student through teacher-trajectory initialization and asymmetric DMD (Yin et al., 2025). Self Forcing subsequently moves distribution matching onto cached autoregressive student rollouts, reducing the mismatch between training and inference histories (Huang et al., 2025). Causal Forcing uses ODE trajectories from an autoregressive teacher for causally compatible initialization, while Causal Forcing++ and Causal-rCM develop more scalable consistency-based initialization and refinement pipelines (Zhu et al., 2026; Zhao et al., 2026; Zheng et al., 2026). Complementary work explores unified autoregressive–bidirectional generation, long-horizon rollout, context management, or future-aware causal supervision (Ma et al., 2026; Liu et al., 2026; Cui et al., 2026; Yang et al., 2026; Chen et al., 2026; Yu et al., 2026; Xu et al., 2026; Zhuang et al., 2026). Recent work on causal video distillation and distribution matching further regularizes the composition of denoising updates, balances mode covering and mode seeking, preserves motion diversity, or aligns teacher supervision with the student's causal information set (Ge et al., 2026; Li et al., 2026; Zou et al., 2026; Bandyopadhyay et al., 2026).

Distribution matching for diffusion distillation. Few-step diffusion distillation includes progressive distillation (Salimans & Ho, 2022), consistency and trajectory matching (Song et al., 2023; Luo et al., 2023a; Zheng et al., 2024), score-based distribution transfer (Luo et al., 2023b; Zhou et al., 2024; Yin et al., 2024b), and adversarial or hybrid objectives (Sauer et al., 2024). DMD updates a generator using the discrepancy between a frozen target score and an online estimate of the evolving generator-distribution score (Yin et al., 2024b). DMD2 improves online score estimation, removes the need for offline regression pairs, incorporates adversarial training, and extends distribution matching to multistep generation (Yin et al., 2024a). Later variants address the mode-seeking behavior and costly online score tracking of DMD through adversarial distribution matching and score-gradient matching objectives (Lu et al., 2025; Wu et al., 2026). Our work builds on asymmetric DMD over causal self-rollouts and studies how teacher–student differences in text-induced score responses can provide directional information for its guided distribution-matching update.

Conditional guidance in diffusion distillation. Classifier-free guidance (CFG) amplifies conditioning through the difference between conditional and unconditional diffusion predictions (Ho & Salimans, 2022). Guidance-distillation methods amortize this composed prediction into a more efficient model through full-model, modular, adapter-based, or embedding-level mechanisms (Meng et al., 2023; Hsiao et al., 2024; Jensen & Sadat, 2025; Fu et al., 2025). SiD-LSG further studies how CFG applied to both pretrained and fake score networks affects score-based one-step distillation (Zhou et al., 2025), while training-free methods adapt the guidance strength or sampling dynamics at inference time (Shen et al., 2024; Gao et al., 2026; Wang et al., 2026). NRP-DMD instead retains the classifier-free-guided teacher target of asymmetric DMD and uses teacher–student disagreement in text-induced score responses to inform the existing distribution-matching update, without introducing a separate response-matching objective or a new inference-time guidance procedure.

3 METHOD

We formulate NRP-DMD as a modification of asymmetric DMD on causal student self-rollouts, as illustrated in Figure 2. NRP-DMD retains the baseline causal rollout construction and CFG-guided teacher target, while introducing a null-relative, blockwise preconditioner for the score discrepancy used to update the generator.

3.1 ASYMMETRIC DMD ON CAUSAL SELF-ROLLOUTS

We consider a few-step causal video generator G_θ refined using asymmetric DMD on student self-rollouts. Given a text condition c and initial Gaussian noise $\boldsymbol{z}$, the generator produces B causal blocks autoregressively:

$$\boldsymbol{x}_0^{1:B} = G_\theta(\boldsymbol{z}, c), \qquad p_\theta(\boldsymbol{x}_0^{1:B} \mid c) = \prod_{b=1}^{B} p_\theta(\boldsymbol{x}_0^b \mid \boldsymbol{x}_0^{<b}, c). \tag{1}$$

Here, $\boldsymbol{x}_0^{1:B} = (\boldsymbol{x}_0^1, \ldots, \boldsymbol{x}_0^B)$, and $\boldsymbol{x}_0$ denotes the concatenated video latent. Each prefix $\boldsymbol{x}_0^{<b}$ is generated by the current student rather than taken from a ground-truth video. The resulting training trajectories therefore follow the same causal rollout process encountered at inference time.

At each training iteration, we sample a diffusion timestep t and perturb the generated trajectory as

$$\boldsymbol{x}_t = \alpha_t \boldsymbol{x}_0 + \sigma_t \boldsymbol{\epsilon}, \qquad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I}). \tag{2}$$

An online fake-score model estimates the score of the evolving student-rollout distribution at $(\boldsymbol{x}_t, t)$, while the frozen bidirectional teacher provides the target score at the same point. Under the target condition c, the online model predicts $\boldsymbol{s}^c_{\text{fake},t} = \boldsymbol{s}_{\text{fake},\psi}(\boldsymbol{x}_t, t \mid c)$. Let $\boldsymbol{s}^c_{\text{real},t}$ and $\boldsymbol{s}^{c^-}_{\text{real},t}$ denote the unguided teacher scores under the target and negative prompts, respectively. Following the asymmetric-DMD baseline, the guided teacher target is

$$\boldsymbol{s}^{\text{cfg},c}_{\text{real},t} = \boldsymbol{s}^{c^-}_{\text{real},t} + \gamma \left(\boldsymbol{s}^c_{\text{real},t} - \boldsymbol{s}^{c^-}_{\text{real},t} \right). \tag{3}$$

Here, c^- denotes the fixed negative prompt used to construct the baseline CFG target and is distinct from the null condition $\varnothing$ introduced below for defining text-induced score responses.

The base DMD score discrepancy is

$$\Delta \boldsymbol{s}_t = \boldsymbol{s}^c_{\text{fake},t} - \boldsymbol{s}^{\text{cfg},c}_{\text{real},t}. \tag{4}$$

This discrepancy serves as the base distribution-matching signal for the generator update.

$$\nabla_\theta \mathcal{L}_{\text{DMD}} = \mathbb{E}_{\boldsymbol{z},c,t,\boldsymbol{\epsilon}} \left[w(t) \left(\frac{\partial \boldsymbol{x}_0}{\partial \theta} \right)^\top \Delta \boldsymbol{s}_t \right], \tag{5}$$

where $w(t)$ absorbs the timestep-dependent weighting and score-parameterization factors. The score discrepancy is treated as stop-gradient when updating G_θ. The online fake-score model is trained to track the evolving student-rollout distribution, whereas the bidirectional teacher remains frozen. NRP-DMD retains this base DMD construction and preconditions $\Delta \boldsymbol{s}_t$ before it enters the generator update.

3.2 NULL-RELATIVE SCORE-RESPONSE DISAGREEMENT

To isolate how conditioning changes each score field at a fixed evaluation point, we evaluate the online fake-score model and the frozen teacher under the null condition $\varnothing$ at the same noised student rollout $\boldsymbol{x}_t$ and diffusion timestep t used for their target-conditioned evaluations. Here, $\varnothing$ denotes the null text condition and is distinct from the negative prompt c^- used to construct the CFG teacher target. We denote the corresponding null scores by

$$\boldsymbol{s}^{\varnothing}_{\text{fake},t} = \boldsymbol{s}_{\text{fake},\psi}(\boldsymbol{x}_t, t \mid \varnothing), \qquad \boldsymbol{s}^{\varnothing}_{\text{real},t} = \boldsymbol{s}_{\text{real}}(\boldsymbol{x}_t, t \mid \varnothing). \tag{6}$$

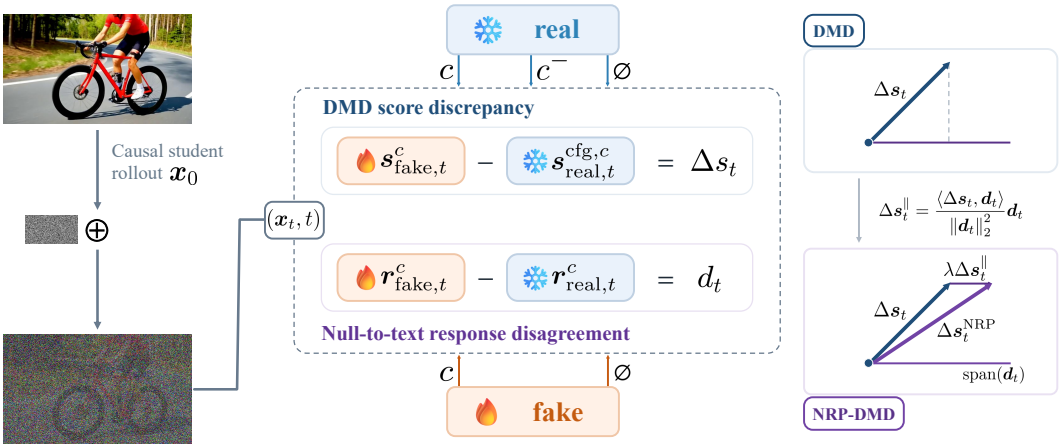

Figure 2: **Overview of NRP-DMD.** A causal student self-rollout is noised and evaluated by an online fake-score model and a frozen bidirectional teacher at the same noisy trajectory and diffusion timestep. The target-conditioned fake score and CFG-guided teacher score define the base DMD signal Δs_t. Separately, the null-to-text responses of the online fake-score model and the unguided teacher define d_t, whose direction is used blockwise to amplify the response-aligned component of the base signal while leaving its orthogonal component unchanged.

For each score model, we define its text-induced score response as the change from the null condition to the target text:

$$r^c_{\text{fake},t} = s^c_{\text{fake},t} - s^{\varnothing}_{\text{fake},t}, \qquad r^c_{\text{real},t} = s^c_{\text{real},t} - s^{\varnothing}_{\text{real},t}. \tag{7}$$

Their difference defines the score-response disagreement

$$\begin{aligned} d_t &= r^c_{\text{fake},t} - r^c_{\text{real},t} \\ &= \left(s^c_{\text{fake},t} - s^{\varnothing}_{\text{fake},t}\right) - \left(s^c_{\text{real},t} - s^{\varnothing}_{\text{real},t}\right) \\ &= \left(s^c_{\text{fake},t} - s^c_{\text{real},t}\right) - \left(s^{\varnothing}_{\text{fake},t} - s^{\varnothing}_{\text{real},t}\right). \end{aligned} \tag{8}$$

The first two lines express d_t as the difference between the two models' text-induced score responses. Equivalently, the final line measures how the unguided fake–teacher score discrepancy changes between the null and target conditions.

For the teacher response, we deliberately use the unguided branches $s^c_{\text{real},t}$ and $s^{\varnothing}_{\text{real},t}$ rather than the CFG-composed target $s^{\text{cfg},c}_{\text{real},t}$. This isolates the teacher's null-to-text response from CFG extrapolation and makes d_t independent of the external guidance scale. The CFG-composed teacher target remains unchanged in the base DMD signal Δs_t.

The null branch of the online fake-score model is trained jointly with its conditional branch through standard condition dropout. When updating the online model with its denoising objective, we replace c by $\varnothing$ with probability ρ, while keeping the noised student rollout and denoising target unchanged. Under the standard denoising-score interpretation, marginalizing over the training prompts makes the dropped-condition branch an estimator of the prompt-marginal student-rollout score. It therefore provides $s^{\varnothing}_{\text{fake},t}$ in Eq. (8) without requiring a separate null-conditioned generator rollout.

3.3 Response-Disagreement Preconditioning

The guided DMD discrepancy Δs_t remains the base distribution-matching signal, while the response-disagreement vector d_t in Eq. (8) provides a sample-dependent reference axis along which the gain of this signal is modulated. NRP-DMD increases the gain of Δs_t along the span of d_t while leaving its orthogonal component unchanged.

For clarity, we suppress the causal block index below. The following operations are performed independently for each sample and causal block, with the inner product and norm taken over all latent dimensions within that block. For $\boldsymbol{d}_t \neq \mathbf{0}$, the component of $\Delta \boldsymbol{s}_t$ parallel to the response-disagreement axis is

$$\Delta \boldsymbol{s}_t^{\parallel} = \frac{\langle \Delta \boldsymbol{s}_t, \boldsymbol{d}_t \rangle}{\|\boldsymbol{d}_t\|_2^2} \boldsymbol{d}_t. \tag{9}$$

We define $\Delta \boldsymbol{s}_t^{\parallel} = \mathbf{0}$ in the degenerate case $\boldsymbol{d}_t = \mathbf{0}$. This normalization makes the projection invariant to any nonzero rescaling of $\boldsymbol{d}_t$, so it depends only on the direction of the response disagreement rather than using its magnitude as an additional sample weight.

Using this projection, we define the sample-dependent preconditioner and the resulting NRP-DMD signal as

$$\boldsymbol{P}_t = \boldsymbol{I} + \lambda \frac{\boldsymbol{d}_t \boldsymbol{d}_t^{\top}}{\|\boldsymbol{d}_t\|_2^2}, \qquad \Delta \boldsymbol{s}_t^{\mathrm{NRP}} = \boldsymbol{P}_t \Delta \boldsymbol{s}_t = \Delta \boldsymbol{s}_t + \lambda \Delta \boldsymbol{s}_t^{\parallel}, \tag{10}$$

with the convention $\boldsymbol{P}_t = \boldsymbol{I}$ when $\boldsymbol{d}_t = \mathbf{0}$. Here, $\lambda \geq 0$ controls the additional gain along the response-disagreement direction; $\boldsymbol{P}_t$ is an identity-plus-rank-one preconditioner for $\boldsymbol{d}_t \neq \mathbf{0}$.

Letting $\Delta \boldsymbol{s}_t^{\perp} = \Delta \boldsymbol{s}_t - \Delta \boldsymbol{s}_t^{\parallel}$ gives the equivalent decomposition

$$\Delta \boldsymbol{s}_t^{\mathrm{NRP}} = \Delta \boldsymbol{s}_t^{\perp} + (1 + \lambda) \Delta \boldsymbol{s}_t^{\parallel}. \tag{11}$$

This decomposition makes the effect of the preconditioner explicit. The response-orthogonal component is preserved, while only the response-aligned component is scaled by $1 + \lambda$.

For $\lambda \geq 0$, the eigenvalue of $\boldsymbol{P}_t$ is $1 + \lambda$ along $\mathrm{span}(\boldsymbol{d}_t)$ and 1 on its orthogonal complement. Hence $\boldsymbol{P}_t$ is positive definite and invertible. In particular, the preconditioning preserves the pointwise zeros of the base DMD discrepancy and maintains positive alignment with the base signal, as

$$\begin{aligned} \Delta \boldsymbol{s}_t^{\mathrm{NRP}} = \mathbf{0} \quad &\Longleftrightarrow \quad \Delta \boldsymbol{s}_t = \mathbf{0}, \\ \left\langle \Delta \boldsymbol{s}_t, \Delta \boldsymbol{s}_t^{\mathrm{NRP}} \right\rangle &= \|\Delta \boldsymbol{s}_t\|_2^2 + \lambda \left\| \Delta \boldsymbol{s}_t^{\parallel} \right\|_2^2 \\ &\geq \|\Delta \boldsymbol{s}_t\|_2^2 . \end{aligned} \tag{12}$$

Unlike directly adding $\lambda \boldsymbol{d}_t$, this preconditioning introduces no correction where the base DMD discrepancy vanishes.

Substituting Eq. (10) into the generator update gives

$$\nabla_\theta \mathcal{L}_{\mathrm{NRP}} = \mathbb{E}_{\boldsymbol{z}, c, t, \boldsymbol{\epsilon}} \left[w(t) \left(\frac{\partial \boldsymbol{x}_0}{\partial \theta} \right)^{\top} \Delta \boldsymbol{s}_t^{\mathrm{NRP}} \right]. \tag{13}$$

As in the base DMD update, all score-derived quantities in Eq. (13) are treated as stop-gradient. The additional null-conditioned score evaluations are used only during distillation and require no changes to the causal generator architecture or its inference procedure.

4 Experiments

4.1 Experimental Setup

Implementation details. We apply NRP-DMD to Self Forcing (Huang et al., 2025) and Causal Forcing (Zhu et al., 2026), using Wan2.1-T2V-1.3B (Wan Team et al., 2025) as the causal generator and Wan2.1-T2V-14B as the frozen real-score model. Within each baseline–NRP-DMD pair, initialization, training data, optimization, rollout, and sampling settings are held fixed. We use four-step sampling, $\lambda = 0.2$, and condition-dropout probability $\rho = 0.1$. Further implementation and training details are provided in Appendix A.

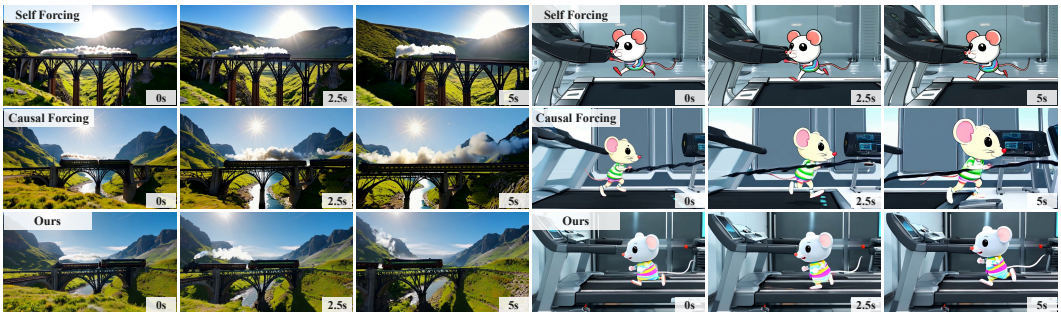

Figure 3: **Qualitative comparison on five-second video generation.** We compare Self Forcing, Causal Forcing, and Ours using the same prompt and random seed. Ours denotes NRP-DMD applied to Causal Forcing.

Table 1: **Five-second video generation.** Model scale, resolution, inference efficiency, and aggregate VBench scores. Throughput (FPS) and first-frame latency (seconds) are measured on a single NVIDIA H100 GPU. * denotes evaluation of the released official Causal Forcing checkpoint.

Model	Params	Resolution	Throughput ↑	Latency ↓	Total ↑	Quality ↑	Semantic ↑
Bidirectional Video Diffusion Models							
LTX-Video (HaCohen et al., 2024)	1.9B	768×512	8.98	13.5	79.83	81.88	71.62
Wan2.1 (Wan Team et al., 2025)	1.3B	832×480	0.78	103	83.37	84.30	79.65
Autoregressive Video Diffusion Models							
NOVA (Deng et al., 2025)	0.6B	768×480	0.88	4.1	80.31	80.66	78.92
Pyramid Flow (Jin et al., 2025)	2B	640×384	6.70	2.5	80.75	83.41	70.11
SkyReels-V2 (Chen et al., 2025)	1.3B	960×540	0.49	112	81.97	83.96	74.01
MAGI-1 (Sand.ai et al., 2025)	4.5B	832×480	0.19	282	78.88	81.67	67.72
Distilled Autoregressive Video Models							
CausVid (Yin et al., 2025)	1.3B	832×480	**17.0**	**0.69**	81.33	83.98	70.72
Self Forcing (Huang et al., 2025)	1.3B	832×480	**17.0**	**0.69**	83.83	84.55	**80.96**
NRP-DMD (Self Forcing)	1.3B	832×480	**17.0**	**0.69**	**84.24**	**85.08**	80.89
Causal Forcing* (Zhu et al., 2026)	1.3B	832×480	**17.0**	**0.69**	84.66	85.56	**81.05**
Causal Forcing (Zhu et al., 2026)	1.3B	832×480	**17.0**	**0.69**	84.63	85.62	80.65
NRP-DMD (Causal Forcing)	1.3B	832×480	**17.0**	**0.69**	**84.98**	**86.01**	80.84

Evaluation. For five-second generation, we follow Self Forcing (Huang et al., 2025) and evaluate on VBench (Huang et al., 2024) using the official prompt suite and evaluation code. Each baseline and its NRP-DMD counterpart use identical generation settings and matched random seeds. For 30-second generation, we follow SALT (Ge et al., 2026), generating five videos for each of 128 MovieGenBench (Polyak et al., 2024) prompts with the same Infinity-RoPE (Yesiltepe et al., 2025) configuration, and evaluate with VBench-Long (Huang et al., 2026).

4.2 Main Results

Five-second video generation. Table 1 shows that NRP-DMD improves Total and Quality for both forcing baselines. On Self Forcing, the Total score increases from 83.83 to 84.24 and the Quality score from 84.55 to 85.08, while the Semantic score changes only slightly from 80.96 to 80.89. On Causal Forcing, NRP-DMD improves Total from 84.63 to 84.98, Quality from 85.62 to 86.01, and Semantic from 80.65 to 80.84. Causal Forcing with NRP-DMD achieves the highest Total and Quality scores among the compared methods. As a training-only modification, NRP-DMD preserves the baselines' 17.0 FPS throughput and 0.69 s first-frame latency, while Figure 3 provides matched qualitative comparisons under identical prompts and random seeds.

Thirty-second long-horizon generation. Although both Self Forcing and Causal Forcing are trained in the standard five-second regime, we further evaluate whether the improvements intro-

Table 2: **Thirty-second video generation.** VBench-Long results on 128 MovieGenBench prompts with five generations per prompt. Bold denotes the best result within each forcing group. * denotes the released official Causal Forcing checkpoint.

Model	Total ↑	Quality ↑	Semantic ↑	Subject Cons. ↑	Background Cons. ↑	Temporal Flick. ↑	Motion Smooth. ↑	Dynamic Degree ↑	Aesthetic Qual. ↑	Imaging Qual. ↑	Overall Cons. ↑
Self Forcing (Huang et al., 2025)	78.36	81.97	**63.93**	96.95	96.03	96.53	98.03	**58.84**	57.42	70.20	**23.27**
NRP-DMD (Self Forcing)	**78.63**	**82.48**	63.21	**97.24**	**96.10**	**97.19**	**98.20**	57.33	**58.00**	**70.89**	23.01
Causal Forcing* (Zhu et al., 2026)	78.08	81.95	62.61	95.83	95.51	94.08	96.89	**83.01**	57.12	70.81	22.79
Causal Forcing (Zhu et al., 2026)	78.52	82.44	62.87	**97.01**	96.03	96.03	98.14	60.58	**58.16**	**72.51**	22.89
NRP-DMD (Causal Forcing)	**79.58**	**83.33**	**64.58**	96.94	**96.08**	**96.31**	**98.24**	70.52	58.03	72.39	**23.51**



Figure 4: **Qualitative comparison on 30-second autoregressive video generation.** We compare Self Forcing, Causal Forcing, and Ours using the same prompt, random seed, and Infinity-RoPE configuration. Ours denotes NRP-DMD applied to Causal Forcing.

duced by NRP-DMD persist over substantially longer autoregressive rollouts. Following the long-horizon protocol described above, we evaluate both baselines and their NRP-DMD counterparts on 30-second generation without additional long-video training. Results on VBench-Long are reported in Table 2.

Across both forcing baselines, NRP-DMD improves the Total and Quality scores on VBench-Long. On Self Forcing, Total increases from 78.36 to 78.63 and Quality from 81.97 to 82.48, with a modest decrease in Semantic from 63.93 to 63.21. On Causal Forcing, the gains are larger, with Total increasing from 78.52 to 79.58, Quality from 82.44 to 83.33, and Semantic from 62.87 to 64.58. These results show that the benefits of NRP-DMD persist over substantially longer autoregressive rollouts without additional long-video training. Figure 4 provides matched qualitative comparisons over the 30-second generation horizon.

4.3 Ablation Studies

We conduct all ablations using the official chunk-wise Causal Forcing pipeline. All variants start from the same causal ODE initialization and use identical training and inference configurations. DMD without condition dropout corresponds to the standard refinement stage of Causal Forcing and serves as the controlled baseline. Unless otherwise specified, we use $\lambda = 0.2$ and condition-dropout probability $\rho = 0.1$.

Component design. We first examine how the response discrepancy $\boldsymbol{d}_t$ should be incorporated into the DMD update. Direct Addition treats it as an independent correction, $\Delta \boldsymbol{s}_t^{\text{add}} = \Delta \boldsymbol{s}_t + \lambda \boldsymbol{d}_t$. Projection instead follows Eq. (10) and amplifies the component of the base DMD signal aligned with the direction defined by $\boldsymbol{d}_t$.

Table 3 shows that condition dropout alone does not improve the DMD baseline, and Direct Addition does not improve its Total or Quality scores. Projection performs better than Direct Addition under both dropout settings, with the full NRP-DMD configuration improving Total, Quality, and Semantic by 0.35, 0.39, and 0.19, respectively. These results show that condition dropout is most effective when paired with projection, consistent with its role in providing the online null branch used to construct the response-disagreement direction.

Table 3: **Component ablation of NRP-DMD.** No/Yes indicates condition dropout; parentheses report changes relative to DMD without dropout.

Metric	DMD		Direct Addition		Projection	
	No	Yes	No	Yes	No	Yes
Total	84.63	84.24 (−0.39)	83.85 (−0.78)	84.27 (−0.36)	84.69 (+0.06)	**84.98** (+**0.35**)
Quality	85.62	85.37 (−0.25)	84.86 (−0.76)	85.15 (−0.47)	85.78 (+0.16)	**86.01** (+**0.39**)
Semantic	80.65	79.69 (−0.96)	79.80 (−0.85)	80.75 (+0.10)	80.31 (−0.34)	**80.84** (+**0.19**)

Table 4: **Direction controls and causal-block placement.** Parentheses report changes relative to the DMD baseline. Early, Late, and Full apply response-disagreement preconditioning to blocks 1–3, 5–7, and 1–7, respectively.

Metric	Reference	Direction Controls		Causal-Block Placement		
	DMD	Uniform Gain	Shuffled Direction	Early (1–3)	Late (5–7)	Full (1–7)
Total	84.63	84.17 (−0.46)	84.15 (−0.48)	84.54 (−0.09)	84.82 (+0.19)	**84.98** (+**0.35**)
Quality	85.62	85.07 (−0.55)	84.95 (−0.67)	85.47 (−0.15)	85.74 (+0.12)	**86.01** (+**0.39**)
Semantic	80.65	80.55 (−0.10)	80.96 (+0.31)	80.83 (+0.18)	**81.16** (+**0.51**)	80.84 (+0.19)

Direction controls and causal-block placement. We design two controls to identify the source of improvement. Uniform Gain rescales the base DMD signal to match the mean relative latent-space signal gain of NRP-DMD, testing whether uniform amplification alone accounts for the improvement. Shuffled Direction uses a cross-instance response direction to construct a correction whose block-wise norm and alignment with the current DMD signal match those of NRP-DMD. It therefore preserves the resulting latent-space signal norm while breaking the correspondence between the response direction and the current training instance. This tests whether matching the correction magnitude and its alignment with the current DMD signal is sufficient.

We further apply response-disagreement preconditioning to early blocks (1–3), late blocks (5–7), or the full trajectory (1–7), while keeping the base DMD signal active at every block. Early and Late precondition the same number of blocks.

In Table 4, Uniform Gain does not reproduce the improvement of NRP-DMD, indicating that uniform amplification alone is insufficient in this control. Shuffled Direction likewise does not reproduce the gains in Total or Quality, supporting the relevance of the sample-dependent response direction. Late-block preconditioning is more effective than early-block preconditioning, while applying the preconditioner throughout the rollout performs best. Additional comparisons of alternative sample-dependent preconditioning axes are provided in Appendix C.3.

5 Conclusion

We introduced Null-Relative Response-Preconditioned DMD (NRP-DMD) for few-step causal video distillation. By comparing teacher and estimated student-rollout null-to-text score responses at the same noised causal self-rollout and diffusion timestep, NRP-DMD defines a sample-dependent response-disagreement direction in score space. The guided DMD discrepancy remains the base distribution-matching signal, while NRP-DMD uses the response-disagreement direction to selectively increase the gain of the response-aligned component while leaving the orthogonal component unchanged. Across Self Forcing and Causal Forcing, NRP-DMD improves generation quality over both 5- and 30-second horizons while maintaining strong text alignment and temporal coherence. As a training-only modification, NRP-DMD requires no changes to the causal generator architecture or additional inference-time computation. These results support teacher–student response disagreement as a useful source of directional information for preconditioning the distribution-matching update across different temporal factorizations.

Reproducibility statement

Implementation details and training settings are provided in Appendix A, together with full benchmark results, response-disagreement diagnostics, bootstrap significance tests, hyperparameter sensitivity analyses, training-overhead measurements, limitations, and additional qualitative results in the supplementary appendix.

AI use statement

We used generative AI tools to assist with code development, experimental design, manuscript editing, language refinement, and formatting. All AI-assisted content was reviewed and verified by the authors. We take responsibility for the final content of this work.

References

Hmrishav Bandyopadhyay, Xuanchi Ren, Zijian Huang, Jay Zhangjie Wu, Tianshi Cao, Ruilong Li, Bryan Chu, Sanja Fidler, Yi-Zhe Song, and Zian Wang. Context-matched distillation: Teacher causality for autoregressive video distillation. *arXiv preprint arXiv:2608.13391*, 2026.

Boyuan Chen, Diego Martí Monsó, Yilun Du, Max Simchowitz, Russ Tedrake, and Vincent Sitzmann. Diffusion forcing: Next-token prediction meets full-sequence diffusion. In *Advances in Neural Information Processing Systems*, volume 37, pp. 24081–24125, 2024. doi: 10.52202/079017-0759.

Guibin Chen, Dixuan Lin, Jiangping Yang, Chunze Lin, Junchen Zhu, Mingyuan Fan, Hao Zhang, Sheng Chen, Zheng Chen, Chengcheng Ma, et al. SkyReels-V2: Infinite-length film generative model. *arXiv preprint arXiv:2504.13074*, 2025.

Shuo Chen, Cong Wei, Sun Sun, Ping Nie, Kai Zhou, Ge Zhang, Ming-Hsuan Yang, and Wenhu Chen. Context forcing: Consistent autoregressive video generation with long context. *arXiv preprint arXiv:2602.06028*, 2026.

Justin Cui, Jie Wu, Ming Li, Tao Yang, Xiaojie Li, Rui Wang, Andrew Bai, Yuanhao Ban, and Cho-Jui Hsieh. Self-forcing++: Towards minute-scale high-quality video generation. In *International Conference on Learning Representations*, 2026.

Haoge Deng, Ting Pan, Haiwen Diao, Zhengxiong Luo, Yufeng Cui, Huchuan Lu, Shiguang Shan, Yonggang Qi, and Xinlong Wang. Autoregressive video generation without vector quantization. In *ICLR*, 2025.

Minghao Fu, Guo-Hua Wang, Xiaohao Chen, Qing-Guo Chen, Zhao Xu, Weihua Luo, and Kaifu Zhang. TeEFusion: Blending text embeddings to distill classifier-free guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2025.

Jiayang Gao, Tianyi Zheng, Jiayang Zou, Fengxiang Yang, Shice Liu, Luyao Fan, Zheyu Zhang, Hao Zhang, Jinwei Chen, Peng-Tao Jiang, Bo Li, and Jia Wang. C^2FG: Control classifier-free guidance via score discrepancy analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2026.

Kaifeng Gao, Jiaxin Shi, Hanwang Zhang, Chunping Wang, Jun Xiao, and Long Chen. Ca2-VDM: Efficient Autoregressive Video Diffusion Model with Causal Generation and Cache Sharing. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267, pp. 18550–18565, 2025a. URL https://mlanthology.org/icml/2025/gao2025icml-ca2vdm/.

Yu Gao, Haoyuan Guo, Tuyen Hoang, Weilin Huang, Lu Jiang, Fangyuan Kong, Huixia Li, Jiashi Li, Liang Li, Xiaojie Li, et al. Seedance 1.0: Exploring the boundaries of video generation models. *arXiv preprint arXiv:2506.09113*, 2025b.

Xingtong Ge, Yi Zhang, Yushi Huang, Dailan He, Xiahong Wang, Bingqi Ma, Guanglu Song, Yu Liu, and Jun Zhang. Salt: Self-consistent distribution matching with cache-aware training for fast video generation. *arXiv preprint arXiv:2604.03118*, 2026.

Yoav HaCohen, Nisan Chiprut, Benny Brazowski, Daniel Shalem, Dudu Moshe, Eitan Richardson, Eran Levin, Guy Shiran, Nir Zabari, Ori Gordon, Poriya Panet, Sapir Weissbuch, Victor Kulikov, Yaki Bitterman, Zeev Melumian, and Ofir Bibi. LTX-Video: Realtime video latent diffusion. *arXiv preprint arXiv:2501.00103*, 2024.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

Yi-Ting Hsiao, Siavash Khodadadeh, Kevin Duarte, Wei-An Lin, Hui Qu, Mingi Kwon, and Ratheesh Kalarot. Plug-and-play diffusion distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.

Xun Huang, Zhengqi Li, Guande He, Mingyuan Zhou, and Eli Shechtman. Self forcing: Bridging the train-test gap in autoregressive video diffusion. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-5576.

Ziqi Huang, Yinan He, Jiashuo Yu, Fan Zhang, Chenyang Si, Yuming Jiang, Yuanhan Zhang, Tianxing Wu, Qingyang Jin, Nattapol Chanpaisit, Yaohui Wang, Xinyuan Chen, Limin Wang, Dahua Lin, Yu Qiao, and Ziwei Liu. VBench: Comprehensive benchmark suite for video generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 21807–21818, 2024.

Ziqi Huang, Fan Zhang, Xiaojie Xu, Yinan He, Jiashuo Yu, Ziyue Dong, Qianli Ma, Nattapol Chanpaisit, Chenyang Si, Yuming Jiang, Yaohui Wang, Xinyuan Chen, Ying-Cong Chen, Limin Wang, Dahua Lin, Yu Qiao, and Ziwei Liu. VBench++: Comprehensive and versatile benchmark suite for video generative models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 48(3):3268–3285, 2026. doi: 10.1109/TPAMI.2025.3633890.

Cristian Perez Jensen and Seyedmorteza Sadat. Efficient distillation of classifier-free guidance using adapters. *Transactions on Machine Learning Research*, 2025.

Yang Jin, Zhicheng Sun, Ningyuan Li, Kun Xu, Kun Xu, Hao Jiang, Nan Zhuang, Quzhe Huang, Yang Song, Yadong Mu, and Zhouchen Lin. Pyramidal flow matching for efficient video generative modeling. In *ICLR*, 2025.

Weijie Kong, Qi Tian, Zijian Zhang, Rox Min, Zuozhuo Dai, Jin Zhou, Jiangfeng Xiong, Xin Li, Bo Wu, Jianwei Zhang, et al. Hunyuanvideo: A systematic framework for large video generative models. *arXiv preprint arXiv:2412.03603*, 2024.

Jiaxing Li, Kai Zou, Cindy Zhou, Kaichen Huang, Junyao Gao, Zile Wang, Yang Liu, Bin Liu, Bo An, and Yangguang Li. DistillAlign: Coordinating mode covering and mode seeking in autoregressive video distillation. *arXiv preprint arXiv:2607.26811*, 2026.

Kunhao Liu, Wenbo Hu, Jiale Xu, Ying Shan, and Shijian Lu. Rolling forcing: Autoregressive long video diffusion in real time. In *International Conference on Learning Representations*, 2026.

Yanzuo Lu, Yuxi Ren, Xin Xia, Shanchuan Lin, Xing Wang, Xuefeng Xiao, Andy J. Ma, Xiaohua Xie, and Jian-Huang Lai. Adversarial distribution matching for diffusion distillation towards efficient image and video synthesis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 16818–16829, 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*, 2023a.

Weijian Luo, Tianyang Hu, Shifeng Zhang, Jiacheng Sun, Zhenguo Li, and Zhihua Zhang. Diff-instruct: A universal approach for transferring knowledge from pre-trained diffusion models. In *Advances in Neural Information Processing Systems*, volume 36, 2023b.

Xinyin Ma, Julius Berner, Chao Liu, Arash Vahdat, Weili Nie, and Xinchao Wang. Flexforcing: Towards a unified autoregressive and bidirectional video diffusion model. *arXiv preprint arXiv:2607.03509*, 2026.

Chenlin Meng, Robin Rombach, Ruiqi Gao, Diederik P. Kingma, Stefano Ermon, Jonathan Ho, and Tim Salimans. On distillation of guided diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023.

Adam Polyak, Amit Zohar, Andrew Brown, Andros Tjandra, Animesh Sinha, Ann Lee, Apoorv Vyas, Bowen Shi, Chih-Yao Ma, Ching-Yao Chuang, et al. Movie gen: A cast of media foundation models. *arXiv preprint arXiv:2410.13720*, 2024.

Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In *International Conference on Learning Representations*, 2022.

Sand.ai, Hansi Teng, Hongyu Jia, Lei Sun, Lingzhi Li, Maolin Li, Mingqiu Tang, Shuai Han, Tianning Zhang, W.Q. Zhang, Weifeng Luo, et al. MAGI-1: Autoregressive video generation at scale. *arXiv preprint arXiv:2505.13211*, 2025.

Axel Sauer, Dominik Lorenz, Andreas Blattmann, and Robin Rombach. Adversarial diffusion distillation. In *Computer Vision – ECCV 2024*, pp. 87–103. Springer, 2024.

Dazhong Shen, Guanglu Song, Zeyue Xue, Fu-Yun Wang, and Yu Liu. Rethinking the spatial inconsistency in classifier-free diffusion guidance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 32211–32252. PMLR, 2023.

Wan Team, Ang Wang, Baole Ai, Bin Wen, Chaojie Mao, Chen-Wei Xie, Di Chen, Feiwu Yu, Haiming Zhao, Jianxiao Yang, Jianyuan Zeng, Jiayu Wang, Jingfeng Zhang, Jingren Zhou, et al. Wan: Open and advanced large-scale video generative models. *arXiv preprint arXiv:2503.20314*, 2025.

Hanyang Wang, Yiyang Liu, Jiawei Chi, Fangfu Liu, Ran Xue, and Yueqi Duan. CFG-Ctrl: Control-based classifier-free diffusion guidance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2026.

Wenhao Wang and Yi Yang. VidProM: A million-scale real prompt-gallery dataset for text-to-video diffusion models. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Zhuguanyu Wu, Ruihao Gong, Yang Yong, Yushi Huang, Xiangyu Fan, Lei Yang, Dahua Lin, and Xianglong Liu. SGMD: Score gradient matching distillation for few-step video diffusion distillation. *arXiv preprint arXiv:2605.30116*, 2026.

Gangwei Xu, Qihang Zhang, Jiaming Zhou, Xing Zhu, Yujun Shen, Xin Yang, and Yinghao Xu. Next forcing: Causal world modeling with multi-chunk prediction. *arXiv preprint arXiv:2606.11187*, 2026.

Shuai Yang, Wei Huang, Ruihang Chu, Yicheng Xiao, Yuyang Zhao, Xianbang Wang, Muyang Li, Enze Xie, Ying-Cong Chen, Yao Lu, Song Han, and Yukang Chen. Longlive: Real-time interactive long video generation. In *International Conference on Learning Representations*, 2026.

Hidir Yesiltepe, Tuna Han Salih Meral, Adil Kaan Akan, Kaan Oktay, and Pinar Yanardag. Infinity-RoPE: Action-controllable infinite video generation emerges from autoregressive self-rollout. *arXiv preprint arXiv:2511.20649*, 2025.

Tianwei Yin, Michaël Gharbi, Taesung Park, Richard Zhang, Eli Shechtman, Fredo Durand, and William T. Freeman. Improved distribution matching distillation for fast image synthesis. *Advances in neural information processing systems*, 37:47455–47487, 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6613–6623, 2024b.

Tianwei Yin, Qiang Zhang, Richard Zhang, William T. Freeman, Fredo Durand, Eli Shechtman, and Xun Huang. From slow bidirectional to fast autoregressive video diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025.

Yonghao Yu, Lang Huang, Runyi Li, Zerun Wang, and Toshihiko Yamasaki. Video-mirai: Autoregressive video diffusion models need foresight. *arXiv preprint arXiv:2606.03971*, 2026.

Min Zhao, Hongzhou Zhu, Kaiwen Zheng, Zihan Zhou, Bokai Yan, Xinyuan Li, Xiao Yang, Chongxuan Li, and Jun Zhu. Causal forcing++: Scalable few-step autoregressive diffusion distillation for real-time interactive video generation. *arXiv preprint arXiv:2605.15141*, 2026.

Jianbin Zheng, Minghui Hu, Zhongyi Fan, Chaoyue Wang, Changxing Ding, Dacheng Tao, and Tat-Jen Cham. Trajectory consistency distillation: Improved latent consistency distillation by semi-linear consistency function with trajectory mapping. *arXiv preprint arXiv:2402.19159*, 2024.

Kaiwen Zheng, Guande He, Min Zhao, Jintao Zhang, Huayu Chen, Jianfei Chen, Chen-Hsuan Lin, Ming-Yu Liu, Jun Zhu, and Qianli Ma. Causal-rcm: A unified teacher-forcing and self-forcing open recipe for autoregressive diffusion distillation in streaming video generation and interactive world models. *arXiv preprint arXiv:2606.25473*, 2026.

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 *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62307–62331. PMLR, 2024.

Mingyuan Zhou, Zhendong Wang, Huangjie Zheng, and Hai Huang. Guided score identity distillation for data-free one-step text-to-image generation. In *International Conference on Learning Representations*, 2025.

Hongzhou Zhu, Min Zhao, Guande He, Hang Su, Chongxuan Li, and Jun Zhu. Causal forcing: Autoregressive diffusion distillation done right for high-quality real-time interactive video generation. *arXiv preprint arXiv:2602.02214*, 2026.

Junhao Zhuang, Shiyi Zhang, Yuxuan Bian, Yaowei Li, Yawen Luo, Yijun Liu, Weiyang Jin, Songchun Zhang, Xianglong He, Xuying Zhang, Haoran Li, Haoyang Huang, Zeyue Xue, and Nan Duan. Self gradient forcing: Native long video extrapolation, 2026. URL `https://arxiv.org/abs/2607.20368`.

Kai Zou, Dian Zheng, Hongbo Liu, Tiankai Hang, Bin Liu, and Nenghai Yu. HiAR: Efficient autoregressive long video generation via hierarchical denoising. *arXiv preprint arXiv:2603.08703*, 2026.

A Implementation Details

A.1 Model Configuration and Baseline Integration

We instantiate NRP-DMD on the chunk-wise variants of Self Forcing (Huang et al., 2025) and Causal Forcing (Zhu et al., 2026). For both baselines, the causal generator is based on Wan2.1-T2V-1.3B (Wan Team et al., 2025) and initialized from the corresponding ODE-distilled baseline checkpoint. The online fake-score model is a bidirectional Wan2.1-T2V-1.3B model, while Wan2.1-T2V-14B serves as the frozen real-score teacher. During asymmetric-DMD refinement, the causal generator and online fake-score model are optimized, whereas the real-score teacher, UMT5-XXL text encoder, and Wan VAE remain frozen. NRP-DMD is applied only to the asymmetric-DMD refinement stage; the baseline-specific causal rollout, cache update, initialization, and inference procedures are otherwise unchanged.

Each training rollout contains 21 latent frames with 16 channels and a spatial resolution of 60×104. We group every three consecutive latent frames into one causal block, yielding seven blocks per rollout. The Wan VAE decodes the full latent sequence into 81 RGB frames at 832×480 resolution, spanning five seconds at 16 fps. The same seven-block partition is used for the causal rollout and for the blockwise NRP-DMD preconditioning.

A.2 Training Setup

The asymmetric-DMD refinement stage uses text prompts sampled from the filtered and LLM-extended VidProM corpus (Wang & Yang, 2024), containing 248,221 prompts in our training split. No real videos or precomputed teacher trajectories are used during refinement. Training is performed on eight NVIDIA A800 GPUs with a global batch size of eight. The causal generator and online fake-score model are optimized with AdamW using learning rates of 2×10^{-6} and 4×10^{-7}, respectively. The fake-score model is updated five times per generator update, following the forcing baselines. We use bfloat16 mixed-precision training with hybrid-sharded FSDP and gradient checkpointing. The gradient norm of each trainable network is clipped to 10. An exponential moving average of the causal generator with decay 0.99 is maintained from training step 200 onward and used for evaluation. All remaining training settings follow the corresponding baseline.

A.3 Sampling and Evaluation Protocols

Model selection. We use a fixed development prompt set, disjoint from the benchmark evaluation prompts, for model selection. Within each baseline–NRP-DMD comparison, we apply the same model-selection protocol and use five-second VBench Total on the development set as the selection criterion. The selected models are then evaluated on the full five-second benchmark and used unchanged for the thirty-second evaluation.

Five-second generation. All paired baseline–NRP-DMD short-video experiments use the same four-step chunk-wise sampler with the nominal denoising schedule $(1000, 750, 500, 250)$, flow shift 5, zero context noise, and three latent frames per causal block. After each block is denoised, a zero-timestep forward pass refreshes the causal KV cache using the clean block. Sampling is performed with batch size one using the EMA generator. The distilled causal generator is directly text conditional and does not use additional classifier-free guidance at inference time.

We evaluate 81-frame, 832×480, 16-fps outputs using the official VBench (Huang et al., 2024) prompt suite and evaluation code. Within every baseline–NRP-DMD pair, prompt preprocessing, sampling settings, and generation seeds are kept identical. Each prompt is evaluated using five generated samples, except for the temporal-flickering subset, for which we follow the official protocol and use 25 samples for each of its 75 prompts. External methods are evaluated using their official checkpoints and recommended inference configurations.

Thirty-second generation. Long-horizon generation requires no additional training. We apply the same Infinity-RoPE (Yesiltepe et al., 2025) configuration to each baseline and its NRP-DMD counterpart. We autoregressively generate 120 latent frames, which decode to 477 RGB frames,

corresponding to approximately 30 seconds at 16 fps. Generation uses three latent frames per causal block, a six-latent-frame local-attention window, and one persistent sink frame.

For evaluation, we use a fixed subset consisting of the first 128 prompts from the released MovieGen-Bench (Polyak et al., 2024) prompt list and generate five matched samples per prompt, yielding 640 videos per method. We evaluate the generated videos using the released VBench-Long implementation from VBench++ (Huang et al., 2026). Following the same evaluation configuration for all methods, each video is uniformly divided into two-second clips and decoded at 8 fps for metric computation, without semantic scene splitting. We report Subject Consistency, Background Consistency, Temporal Flickering, Motion Smoothness, Dynamic Degree, Aesthetic Quality, Imaging Quality, and Overall Consistency. Temporal Flickering is computed without RAFT-based static filtering. All methods use identical prompts, matched sample assignments, video preprocessing, and evaluation settings.

Efficiency evaluation. Inference throughput and first-frame latency are measured on a single NVIDIA H100 GPU with batch size one and bfloat16 precision, following the timing protocol of the corresponding forcing baseline. Paired quantitative and qualitative comparisons use identical generation seeds.

A.4 Implementation of Controlled Variants

All controlled variants use the same Causal Forcing initialization, training configuration, prompt stream, and inference protocol as the corresponding NRP-DMD setup. For the ablation study, we use a fixed training seed for all variants and do not perform separate model selection. We use $\lambda = 0.2$ by default. The No and Yes condition-dropout settings use $\rho = 0$ and $\rho = 0.1$, respectively, while leaving all other settings unchanged.

Direct Addition. Direct Addition treats the response disagreement as an additive correction rather than using it to define a projection axis

$$\Delta \boldsymbol{s}_t^{\mathrm{add}} = \Delta \boldsymbol{s}_t + \lambda \boldsymbol{d}_t. \tag{14}$$

Uniform Gain. Uniform Gain uniformly rescales the base DMD signal as

$$\Delta \boldsymbol{s}_t^{\mathrm{uni}} = \bar{r}\, \Delta \boldsymbol{s}_t, \qquad \bar{r} = 1.025. \tag{15}$$

We determine $\bar{r}$ from the corresponding NRP-DMD training run as the empirical mean of

$$\frac{\left\|\Delta \boldsymbol{s}_t^{\mathrm{NRP}}\right\|_2}{\left\|\Delta \boldsymbol{s}_t\right\|_2}, \tag{16}$$

where both norms are taken over the full latent update tensor at each generator update, and the average is computed over the training updates prior to benchmark evaluation. This control therefore matches the mean pre-optimizer signal gain of NRP-DMD, but not its sample- and block-dependent preconditioning.

Shuffled Direction. Shuffled Direction preserves the magnitude and alignment of the NRP-DMD response-aligned component while breaking its correspondence with the current training sample. For each sample, we cyclically take the response-disagreement direction from the next distributed training sample and denote the resulting direction for block b by $\boldsymbol{d}_{t,b}^{\mathrm{shuf}}$.

Following the notation of the main paper, the response-aligned component used by NRP-DMD at block b is

$$\Delta \boldsymbol{s}_{t,b}^{\parallel} = \frac{\langle \Delta \boldsymbol{s}_{t,b}, \boldsymbol{d}_{t,b} \rangle}{\left\|\boldsymbol{d}_{t,b}\right\|_2^2} \boldsymbol{d}_{t,b}. \tag{17}$$

We first remove from $\boldsymbol{d}_{t,b}^{\mathrm{shuf}}$ its component parallel to the current base DMD signal and normalize the remaining direction:

$$\boldsymbol{u}_{t,b} = \frac{\boldsymbol{d}_{t,b}^{\mathrm{shuf}} - \frac{\langle \boldsymbol{d}_{t,b}^{\mathrm{shuf}}, \Delta \boldsymbol{s}_{t,b} \rangle}{\left\|\Delta \boldsymbol{s}_{t,b}\right\|_2^2} \Delta \boldsymbol{s}_{t,b}}{\left\| \boldsymbol{d}_{t,b}^{\mathrm{shuf}} - \frac{\langle \boldsymbol{d}_{t,b}^{\mathrm{shuf}}, \Delta \boldsymbol{s}_{t,b} \rangle}{\left\|\Delta \boldsymbol{s}_{t,b}\right\|_2^2} \Delta \boldsymbol{s}_{t,b} \right\|_2}. \tag{18}$$

Table 5: **Per-dimension Quality scores on five-second VBench.** We report the complete Quality breakdown for Causal Forcing and its NRP-DMD counterpart. Bold denotes the better result within the pair.

Method	Subject Consistency	Background Consistency	Temporal Flickering	Motion Smoothness	Dynamic Degree	Aesthetic Quality	Imaging Quality	Quality Score
Causal Forcing (Zhu et al., 2026)	96.54	**96.62**	**99.30**	98.55	65.83	**67.30**	70.96	85.62
NRP-DMD (Causal Forcing)	**96.60**	96.09	99.00	**98.75**	**73.89**	66.43	**71.10**	**86.01**

Table 6: **Per-dimension Semantic scores on five-second VBench.** We report the complete Semantic breakdown for Causal Forcing and its NRP-DMD counterpart. Bold denotes the better result within the pair.

Method	Object Class	Multiple Objects	Human Action	Color	Spatial Relation	Scene	Temporal Style	Appearance Style	Overall Consistency	Semantic Score
Causal Forcing (Zhu et al., 2026)	**96.36**	87.04	94.80	**86.78**	80.06	**56.99**	**24.34**	**20.48**	26.56	80.65
NRP-DMD (Causal Forcing)	95.03	**87.90**	**96.80**	85.42	**81.45**	56.92	24.27	20.42	**26.82**	**80.84**

We then define the cosine between the NRP-DMD response-aligned component and the base DMD signal as

$$c_{t,b} = \frac{\left\langle \Delta \boldsymbol{s}_{t,b}, \Delta \boldsymbol{s}^{\parallel}_{t,b} \right\rangle}{\left\| \Delta \boldsymbol{s}_{t,b} \right\|_2 \left\| \Delta \boldsymbol{s}^{\parallel}_{t,b} \right\|_2}. \tag{19}$$

Using the shuffled orthogonal direction, we construct

$$\Delta \boldsymbol{s}^{\mathrm{shuf},\parallel}_{t,b} = \left\| \Delta \boldsymbol{s}^{\parallel}_{t,b} \right\|_2 \left(c_{t,b} \frac{\Delta \boldsymbol{s}_{t,b}}{\left\| \Delta \boldsymbol{s}_{t,b} \right\|_2} + \sqrt{1 - c^2_{t,b}}\, \boldsymbol{u}_{t,b} \right). \tag{20}$$

The resulting control signal is

$$\Delta \boldsymbol{s}^{\mathrm{shuf}}_{t,b} = \Delta \boldsymbol{s}_{t,b} + \lambda \Delta \boldsymbol{s}^{\mathrm{shuf},\parallel}_{t,b}. \tag{21}$$

By construction,

$$\left\| \Delta \boldsymbol{s}^{\mathrm{shuf},\parallel}_{t,b} \right\|_2 = \left\| \Delta \boldsymbol{s}^{\parallel}_{t,b} \right\|_2, \qquad \cos\left(\Delta \boldsymbol{s}_{t,b}, \Delta \boldsymbol{s}^{\mathrm{shuf},\parallel}_{t,b} \right) = \cos\left(\Delta \boldsymbol{s}_{t,b}, \Delta \boldsymbol{s}^{\parallel}_{t,b} \right). \tag{22}$$

Thus, Shuffled Direction matches the blockwise magnitude and alignment of the NRP-DMD response-aligned component, and consequently the norm of the resulting preconditioned signal, while breaking the correspondence between the response-disagreement direction and the current prompt–sample pair.

Causal-block placement. For the causal-block placement controls, the base DMD update remains active at all seven causal blocks, while response-disagreement preconditioning is applied only to the specified subset. Early applies NRP-DMD to blocks 1–3, Late to blocks 5–7, and Full to blocks 1–7. Early and Late therefore precondition the same number of causal blocks, while Full corresponds to the default NRP-DMD configuration.

B Per-Dimension Five-Second VBench Results

We provide the complete per-dimension VBench results for the five-second Causal Forcing comparison in Tables 5 and 6. The former reports the dimensions contributing to the VBench Quality score, while the latter reports the dimensions contributing to the Semantic score. Causal Forcing and its NRP-DMD counterpart use the same prompts, generation settings, and matched random seeds as in the main evaluation.

C Response-Disagreement Interpretation and Diagnostics

C.1 Conditional-Residual and Sensitivity Interpretation

We first provide an interpretation of the response-disagreement vector defined in Eq. (8). At a fixed noised rollout $\boldsymbol{x}_t$ and diffusion timestep t, define the unguided fake–teacher score residual under condition c as

$$\boldsymbol{e}_t(c) = \boldsymbol{s}^c_{\mathrm{fake},t} - \boldsymbol{s}^c_{\mathrm{real},t}. \tag{23}$$

Then the response disagreement can be written exactly as

$$\boldsymbol{d}_t = \boldsymbol{e}_t(c) - \boldsymbol{e}_t(\varnothing). \tag{24}$$

Thus, $\boldsymbol{d}_t$ removes the fake–teacher residual already present at the null reference and isolates how this residual changes when moving from the null condition to the target text. In this sense, $\boldsymbol{d}_t$ captures the condition-dependent component of the mismatch between the estimated student-rollout score field and the unguided teacher score field, rather than their absolute mismatch.

To make this interpretation more explicit, let $\boldsymbol{h}(c)$ denote the text-conditioning representation supplied to the score models, with $\boldsymbol{h}_c = \boldsymbol{h}(c)$, $\boldsymbol{h}_\varnothing = \boldsymbol{h}(\varnothing)$, and $\delta\boldsymbol{h} = \boldsymbol{h}_c - \boldsymbol{h}_\varnothing$. For analytical purposes, consider the straight-line path in conditioning-representation space

$$\boldsymbol{h}_\tau = \boldsymbol{h}_\varnothing + \tau\delta\boldsymbol{h}, \qquad \tau \in [0,1]. \tag{25}$$

Writing the unguided residual as $\boldsymbol{e}_t(\boldsymbol{h})$, if the score fields are differentiable with respect to the conditioning representation along this path, the fundamental theorem of calculus gives

$$\begin{aligned}
\boldsymbol{d}_t &= \boldsymbol{e}_t(\boldsymbol{h}_c) - \boldsymbol{e}_t(\boldsymbol{h}_\varnothing) && (26)\\
&= \int_0^1 \boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{e}_t(\boldsymbol{h}_\tau)\,\delta\boldsymbol{h}\,d\tau && (27)\\
&= \int_0^1 \left[\boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{s}_{\mathrm{fake},t}(\boldsymbol{h}_\tau) - \boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{s}_{\mathrm{real},t}(\boldsymbol{h}_\tau)\right]\delta\boldsymbol{h}\,d\tau. && (28)
\end{aligned}$$

Therefore, $\boldsymbol{d}_t$ can be interpreted as the integrated mismatch between how the estimated student-rollout score field and the teacher score field respond to the change from the null condition to the target text. Under a local linearization around the null condition,

$$\boldsymbol{d}_t \approx \left[\boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{s}_{\mathrm{fake},t}(\boldsymbol{h}_\varnothing) - \boldsymbol{J}_{\boldsymbol{h}}\boldsymbol{s}_{\mathrm{real},t}(\boldsymbol{h}_\varnothing)\right]\delta\boldsymbol{h}. \tag{29}$$

This gives $\boldsymbol{d}_t$ a conditional-sensitivity interpretation: it measures, along the null-to-text conditioning path, how differently the estimated student-rollout and teacher score fields respond to the same conditioning change, while the noised rollout and diffusion timestep $(\boldsymbol{x}_t, t)$ remain fixed.

C.2 Response Disagreement as a Preconditioning Axis

The analysis above identifies $\boldsymbol{d}_t$ as a condition-specific, sample-dependent direction, but does not by itself prescribe how this information should modify the generator update. Motivated by this interpretation, NRP-DMD uses $\boldsymbol{d}_t$ only to define where the existing DMD signal receives additional gain, rather than treating $\boldsymbol{d}_t$ itself as a new correction.

As in Section 3.3, we suppress the causal-block index for clarity; all operations below are applied independently to each sample and causal block.

For $\boldsymbol{d}_t \neq 0$, let

$$\boldsymbol{u}_t = \frac{\boldsymbol{d}_t}{\|\boldsymbol{d}_t\|_2}. \tag{30}$$

The NRP-DMD preconditioner can then be written as

$$\boldsymbol{P}_t = \boldsymbol{I} + \lambda\boldsymbol{u}_t\boldsymbol{u}_t^\top. \tag{31}$$

It is a symmetric linear transformation that leaves every direction orthogonal to $\boldsymbol{d}_t$ unchanged and scales the component along $\mathrm{span}(\boldsymbol{d}_t)$ by $1+\lambda$.

The corresponding quadratic form makes this directional reweighting explicit:

$$\Delta \boldsymbol{s}_t^\top \boldsymbol{P}_t \Delta \boldsymbol{s}_t = \|\Delta \boldsymbol{s}_t\|_2^2 + \lambda \left(\boldsymbol{u}_t^\top \Delta \boldsymbol{s}_t\right)^2. \tag{32}$$

Thus, at the level of the score-space signal geometry, NRP-DMD assigns additional weight only to the component of the existing DMD discrepancy that overlaps the response-disagreement direction. This provides an anisotropic reweighting interpretation of the preconditioned signal, rather than defining a separate global scalar objective for NRP-DMD.

This interpretation also distinguishes preconditioning from directly adding the response disagreement. Direct Addition uses

$$\Delta \boldsymbol{s}_t + \lambda \boldsymbol{d}_t, \tag{33}$$

so $\boldsymbol{d}_t$ itself becomes an independent correction whose magnitude affects the update. In contrast, NRP-DMD is invariant to nonzero rescaling of $\boldsymbol{d}_t$ and modifies only the part of the base DMD signal already aligned with its span. In particular,

$$\Delta \boldsymbol{s}_t = 0 \quad \Longrightarrow \quad \Delta \boldsymbol{s}_t^{\mathrm{NRP}} = 0, \tag{34}$$

whereas a direct additive correction need not vanish.

Accordingly, the response disagreement determines *where* the base distribution-matching signal is emphasized, while the DMD discrepancy continues to determine *what* correction is applied.

The analysis above motivates using response disagreement as a directional axis, but does not by itself determine whether the response-aligned component should be amplified or suppressed, nor the appropriate gain magnitude. We examine this choice empirically. As shown in Appendix E, suppressing the response-aligned component with $\lambda = -0.2$ degrades both Total and Quality, whereas moderate positive preconditioning improves both metrics. Moreover, the Uniform Gain and Shuffled Direction controls in Section 4.3 show that the improvement is not reproduced by uniform amplification of the DMD signal or by a matched but sample-mismatched direction. Together, these results motivate positive, sample-specific amplification as an empirical design choice rather than a consequence of the analytical interpretation alone.

C.3 Preconditioning Axes and Direction Controls

We extend the direction controls in Section 4.3 with two alternative sample-dependent preconditioning axes. Teacher Response uses the teacher's null-to-text score response $\boldsymbol{r}^c_{\mathrm{real},t}$ in Eq. (7), while Target-conditioned Residual uses the unguided fake–teacher score residual $\boldsymbol{e}_t(c)$ in Eq. (23). The former does not compare the teacher and fake-model responses, whereas the latter does not subtract the null-reference residual $\boldsymbol{e}_t(\varnothing)$. Each alternative replaces $\boldsymbol{d}_t$ in the blockwise preconditioner in Eq. (10), retaining the base guided DMD signal and the preconditioning strength $\lambda = 0.2$.

Table 7 shows that neither alternative axis reproduces the Total and Quality gains of NRP-DMD. Among the axes evaluated with $\rho = 0$, Response Disagreement achieves higher Total and Quality than either alternative. These results support using disagreement between the two models' null-to-text responses to define the preconditioning axis, rather than the teacher response alone or the target-conditioned residual.

C.4 Empirical Directional Diagnostics

We next examine whether the response-disagreement direction exhibits instance-specific correspondence with the current DMD signal. We compare its directional overlap with that of a response direction taken from an unrelated training instance, and separately examine whether directional relevance is explained by the magnitude of the response disagreement itself. Unless otherwise specified, all diagnostics are collected from the Causal Forcing NRP-DMD refinement run with $\lambda = 0.2$ and $\rho = 0.1$, using the same noised causal self-rollouts on which the generator update is computed.

Instance-specific directional alignment. For causal block b, we measure the fraction of the base DMD squared norm lying along the response-disagreement axis:

$$A^{\mathrm{self}}_{t,b} = \frac{\langle \Delta \boldsymbol{s}_{t,b}, \boldsymbol{d}_{t,b} \rangle^2}{\|\Delta \boldsymbol{s}_{t,b}\|_2^2 \, \|\boldsymbol{d}_{t,b}\|_2^2}. \tag{35}$$

Table 7: **Alternative preconditioning axes on five-second VBench.** Teacher Response and Target-conditioned Residual are newly evaluated; the remaining results are reproduced from Tables 3 and 4. Response Disagreement at $\rho = 0$ corresponds to the Projection/No setting in Table 3. Bold denotes the best result in each metric column.

Method	ρ	Total ↑	Quality ↑	Semantic ↑
References				
DMD	0	84.63	85.62	80.65
DMD + condition dropout	0.1	84.24	85.37	79.69
Direction Controls				
Uniform Gain	0.1	84.17	85.07	80.55
Shuffled Direction	0.1	84.15	84.95	80.96
Preconditioning Axes				
Teacher Response	0	84.60	85.52	80.94
Target-conditioned Residual	0	84.23	85.03	**81.01**
Response Disagreement	0	84.69	85.78	80.31
NRP-DMD (Causal Forcing)	0.1	**84.98**	**86.01**	80.84

This quantity measures the fraction of the base DMD squared norm contained in its response-aligned component.

To examine whether this overlap is specific to the current training instance, we replace $\boldsymbol{d}_{t,b}$ with a response-disagreement direction from another sample evaluated at the same diffusion timestep:

$$A_{t,b}^{\text{cross}} = \frac{\left\langle \Delta \boldsymbol{s}_{t,b}^{(i)}, \boldsymbol{d}_{t,b}^{(j)} \right\rangle^2}{\left\| \Delta \boldsymbol{s}_{t,b}^{(i)} \right\|_2^2 \left\| \boldsymbol{d}_{t,b}^{(j)} \right\|_2^2}, \qquad i \neq j. \tag{36}$$

We use a response-disagreement direction from another distributed sample as the cross-instance donor. As summarized in Table 8, the current-sample axis captures 11.67% of the DMD squared norm on average, compared with 0.0086% for the cross-instance axis. Moreover, 99.75% of sample–block measurements satisfy $A_{t,b}^{\text{self}} > A_{t,b}^{\text{cross}}$, and the same ordering holds for every prompt after prompt-level aggregation. The response-disagreement axis therefore exhibits a strong instance-specific correspondence with the DMD signal of its own prompt–rollout pair, consistent with its interpretation as a condition-dependent mismatch specific to the current rollout state.

Disagreement magnitude and directional relevance. The magnitude of response disagreement and its directional overlap with the DMD signal characterize different properties. We measure the normalized response-disagreement magnitude as

$$D_{t,b} = \frac{\left\| \boldsymbol{r}_{\text{fake},t,b}^{c} - \boldsymbol{r}_{\text{real},t,b}^{c} \right\|_2}{\left\| \boldsymbol{r}_{\text{fake},t,b}^{c} \right\|_2 + \left\| \boldsymbol{r}_{\text{real},t,b}^{c} \right\|_2}. \tag{37}$$

Table 8 shows that response magnitude and directional overlap exhibit different trends across diffusion timesteps. Toward noisier states, normalized disagreement decreases slightly, whereas the fraction of DMD squared norm aligned with the response-disagreement axis increases. Thus, lower normalized disagreement can coincide with greater directional overlap across the reported timestep groups.

This distinction is consistent with the use of response disagreement as an adaptive axis. Let $\Pi_{\boldsymbol{d}}(\cdot)$ denote orthogonal projection onto $\text{span}(\boldsymbol{d})$. For any nonzero scalar α,

$$\Pi_{\alpha \boldsymbol{d}_{t,b}} \left(\Delta \boldsymbol{s}_{t,b} \right) = \Pi_{\boldsymbol{d}_{t,b}} \left(\Delta \boldsymbol{s}_{t,b} \right), \tag{38}$$

so the NRP-DMD projection depends on the subspace spanned by the response disagreement rather than its scale. The Direct Addition control in the main paper provides the complementary comparison in which the response disagreement is used directly as an additive signal.

Table 8: **Response-disagreement diagnostics across diffusion timesteps.** We report response-aligned energy, cross-response energy, normalized response disagreement, and the resulting NRP-DMD modification over different diffusion timestep ranges.

Timestep	Self energy $A^{\text{self}}_{t,b}$	Cross energy $A^{\text{cross}}_{t,b}$	Disagreement $D_{t,b}$	Added/Base norm $R^{\text{add}}_{t,b}$	NRP/Base norm $R^{\text{NRP}}_{t,b}$
20–249	8.72%	0.0029%	0.703	5.86%	1.019
250–499	9.31%	0.0010%	0.701	6.05%	1.020
500–749	13.58%	0.0026%	0.695	7.29%	1.029
750–980	15.09%	0.0279%	0.647	7.42%	1.032
Overall	**11.67%**	**0.0086%**	**0.687**	**6.66%**	**1.025**

Table 9: **Response-aligned DMD energy across causal blocks.** We report A^{self}_b averaged over diffusion timesteps and diagnostic samples. The response-aligned energy remains nearly uniform across the seven causal blocks.

Block	1	2	3	4	5	6	7
A^{self}_b	11.85%	11.51%	11.58%	11.58%	11.76%	11.77%	11.67%

To quantify the magnitude of the NRP-DMD modification, we define

$$R^{\text{add}}_{t,b} = \frac{\left\| \Delta \boldsymbol{s}^{\text{NRP}}_{t,b} - \Delta \boldsymbol{s}_{t,b} \right\|_2}{\left\| \Delta \boldsymbol{s}_{t,b} \right\|_2} = \frac{\left\| \lambda \Delta \boldsymbol{s}^{\parallel}_{t,b} \right\|_2}{\left\| \Delta \boldsymbol{s}_{t,b} \right\|_2}, \tag{39}$$

and

$$R^{\text{NRP}}_{t,b} = \frac{\left\| \Delta \boldsymbol{s}^{\text{NRP}}_{t,b} \right\|_2}{\left\| \Delta \boldsymbol{s}_{t,b} \right\|_2}. \tag{40}$$

The former measures the relative norm of the added response-aligned component, while the latter measures the norm ratio of the complete preconditioned signal. All quantities are computed per sample–block observation and then averaged within each reported timestep range. The NRP-DMD modification statistics are computed using the default preconditioning strength $\lambda = 0.2$.

Variation across diffusion timesteps and causal blocks. As shown in Table 8, the directional overlap between the response-disagreement axis and the base DMD signal increases at noisier diffusion timesteps. In contrast, Table 9 shows that the response-aligned energy remains similar across the seven causal blocks. Thus, the response-aligned component persists throughout the causal rollout, while its strength varies more substantially with diffusion timestep than with block position. Notably, Early and Late blocks exhibit comparable average response-aligned energy despite their different performance in the block-placement ablation, suggesting that the placement effect is not determined solely by the degree of directional overlap.

Magnitude of the resulting update. NRP-DMD produces a targeted directional modification without substantially rescaling the overall DMD signal. As reported in Table 8, the additional response-aligned component has an average norm equal to 6.66% of the base DMD signal, while the resulting NRP-DMD signal has an average norm ratio of 1.025. Thus, the modification is concentrated along a sample-specific response-disagreement direction rather than expressed as a large change in overall signal magnitude. Together with the Uniform Gain control, these results support the interpretation that NRP-DMD benefits from *where* the base update is amplified, rather than from a broad increase in signal magnitude.

D Bootstrap Confidence Intervals for Main Results

We characterize the statistical uncertainty of the five-second VBench improvements using paired prompt-level bootstrap resampling. The analysis is performed separately for Self Forcing and Causal Forcing, comparing each baseline only with its matched NRP-DMD counterpart.

Table 10: **Paired bootstrap analysis on five-second VBench.** Entries report NRP-DMD minus baseline differences with 95% prompt-level bootstrap confidence intervals from 10,000 replicates. Positive values favor NRP-DMD.

Baseline	Δ Total 95% CI	Δ Quality 95% CI	Δ Semantic 95% CI
Self Forcing	$+0.41$ [0.08, 0.74]	$+0.53$ [0.09, 0.96]	-0.07 [-0.11, -0.03]
Causal Forcing	$+0.35$ [0.12, 0.58]	$+0.39$ [0.15, 0.63]	$+0.19$ [-0.04, 0.42]

For each VBench dimension, we first average the metric value over the matched generation seeds for each prompt and method. We then resample prompts with replacement using the same resampling indices for the baseline and NRP-DMD. Within each bootstrap replicate, the dimension-level scores are recomputed from the resampled prompts, and Quality, Semantic, and Total are subsequently obtained using the official VBench aggregation procedure. This preserves the paired structure of the evaluation rather than treating the final aggregate scores as independent observations.

We use 10,000 bootstrap replicates and report the point estimate of the NRP-DMD minus baseline difference together with its 95% percentile bootstrap confidence interval. Table 10 summarizes the results for the main five-second VBench metrics.

The bootstrap analysis yields positive point estimates for Total and Quality under both forcing baselines, consistent with the aggregate improvements reported in the main paper. Importantly, the 95% confidence intervals for both Total and Quality remain above zero for Self Forcing and Causal Forcing, indicating that these improvements are robust to prompt-level variation under the paired bootstrap analysis. For Semantic, Self Forcing shows a small negative difference of -0.07 with a confidence interval entirely below zero, whereas the positive difference for Causal Forcing has a confidence interval that overlaps zero. Thus, the most consistent gains of NRP-DMD are observed in the Total and Quality metrics, while its effect on Semantic alignment is smaller and baseline-dependent.

E Hyperparameter Sensitivity

We examine the sensitivity of NRP-DMD to the preconditioning strength λ and the condition-dropout probability ρ. All experiments use the Causal Forcing setup from the main ablation study with the same initialization, training configuration, and evaluation protocol. When varying one hyperparameter, the other is fixed to its default value, $\lambda = 0.2$ and $\rho = 0.1$. For diagnostic purposes, we additionally evaluate $\lambda = -0.2$ outside the default amplification regime $\lambda \geq 0$ to test suppression of the response-aligned component, alongside the positive sweep.

The coefficient λ controls the gain applied to the response-aligned component of the base DMD signal. Positive values amplify this component, $\lambda = 0$ removes response-disagreement preconditioning, and negative values suppress it. Because the λ study fixes $\rho = 0.1$, the $\lambda = 0$ setting still trains the online fake-score model with condition dropout and is therefore distinct from the standard DMD baseline with $\rho = 0$. The parameter ρ controls how frequently the online fake-score model is trained under the null condition used to construct the response-disagreement direction. Table 11 reports the five-second VBench results.

The λ study shows a clear preference for moderate positive preconditioning. Suppressing the response-aligned component with $\lambda = -0.2$ substantially reduces both Total and Quality, while the $\lambda = 0$ setting, which retains condition dropout but removes response-disagreement preconditioning, also underperforms the standard DMD reference. This indicates that condition dropout alone does not account for the improvement. Positive preconditioning consistently improves Total and Quality over the dropout-only control, with $\lambda = 0.2$ achieving the best scores on both metrics. Increasing the gain further provides no additional aggregate benefit, although $\lambda = 0.3$ yields the highest Semantic score. Overall, these results support selectively amplifying rather than suppressing the response-aligned component, with moderate amplification providing the best trade-off.

Table 11: **Hyperparameter sensitivity of NRP-DMD.** We vary λ with $\rho = 0.1$, and vary ρ with $\lambda = 0.2$. The standard DMD setting is included as reference. Bold denotes the best result within each sweep.

Setting	Total ↑	Quality ↑	Semantic ↑
Reference			
DMD ($\lambda = 0,\ \rho = 0$)	84.63	85.62	80.65
λ sweep ($\rho = 0.1$)			
$\lambda = -0.2$	83.68	84.42	80.73
$\lambda = 0.0$	84.24	85.37	79.69
$\lambda = 0.1$	84.92	85.96	80.77
$\lambda = 0.2$	**84.98**	**86.01**	80.84
$\lambda = 0.3$	84.83	85.79	**80.96**
$\lambda = 0.5$	84.71	85.82	80.29
ρ sweep ($\lambda = 0.2$)			
$\rho = 0.1$	**84.98**	**86.01**	80.84
$\rho = 0.2$	84.73	85.87	80.18
$\rho = 0.5$	84.32	85.11	**81.16**

For the condition-dropout probability, $\rho = 0.1$ gives the strongest overall performance, achieving the best Total and Quality scores. Larger dropout probabilities reduce these aggregate metrics, although $\rho = 0.5$ yields the highest Semantic score.

F Training-Time Overhead

NRP-DMD introduces only modest additional computation during distillation. During each generator update, it evaluates the online fake-score model and the frozen real-score teacher once more under the null condition at the same noised student rollout and diffusion timestep used by the base DMD update. These additional evaluations reuse the same rollout, noise realization, and timestep and do not require a second causal generator rollout.

Based on the end-to-end refinement logs under the Causal Forcing setup, NRP-DMD increases wall-clock training time by approximately 10% relative to the corresponding asymmetric-DMD baseline. This additional computation is confined to the distillation stage; inference uses the same causal generator architecture and sampling procedure as the baseline and therefore introduces no additional inference-time computation.

G Limitations

Our evaluation focuses on Wan2.1-based Self Forcing and Causal Forcing frameworks, with generation horizons of up to 30 seconds. Further evaluation is needed to assess how well NRP-DMD transfers to other model backbones, different causal distillation frameworks, and substantially longer causal rollouts.

H Additional Qualitative Results

Figures 5 and 6 show additional five- and 30-second comparisons, respectively. Each baseline–NRP-DMD pair uses identical prompts, seeds, and generation settings, including the Infinity-RoPE configuration for 30-second rollouts.

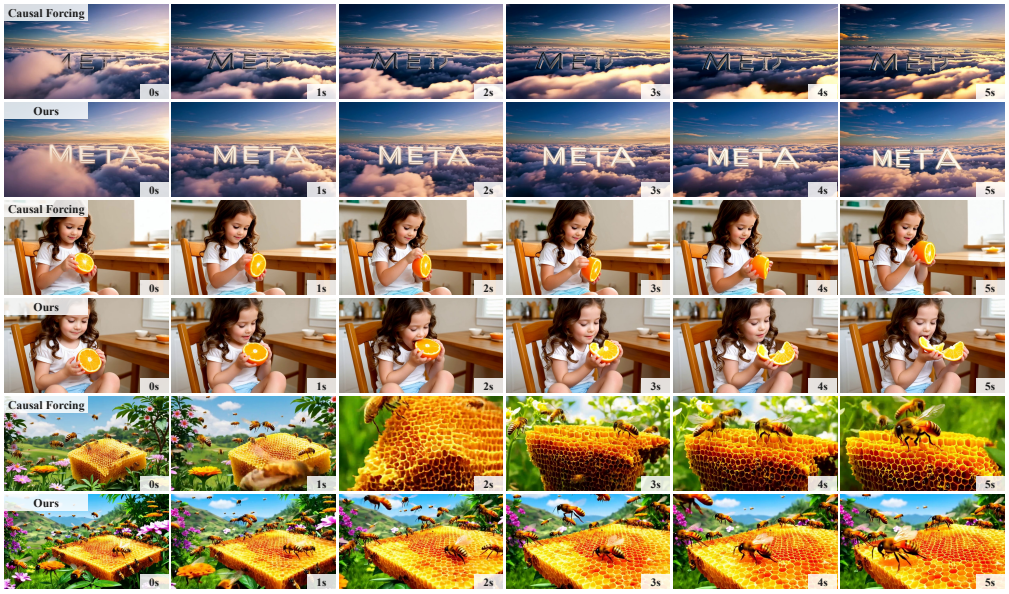

Figure 5: **Five-second qualitative comparisons.** Prompt excerpts (top to bottom): (1) *A dynamic landscape photograph where clouds flow and shift to form the word "Meta." ...* (2) *a young girl sitting on a wooden chair, peeling an orange ...* (3) *A vibrant ecosystem bustling with activity, primarily composed of hundreds of bees. ...*

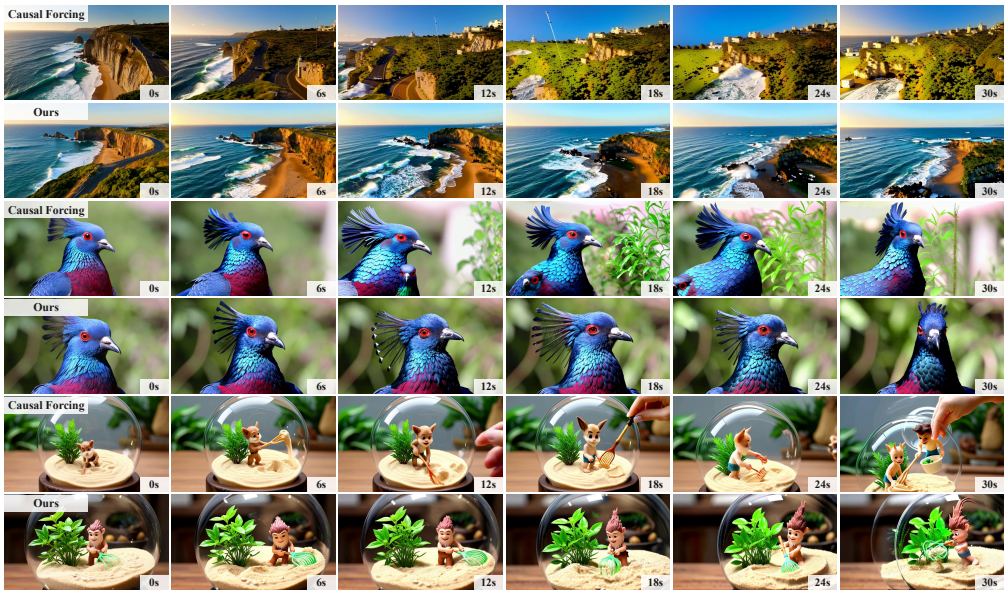

Figure 6: **30-second qualitative comparisons.** Prompt excerpts (top to bottom): (1) *Drone view of waves crashing against the rugged cliffs ...* (2) *This close-up shot of a Victoria crowned pigeon showcases its striking blue plumage and red chest. ...* (3) *A close up view of a glass sphere that has a zen garden within it. ...*