PHYSGUIDE: PHYSICS-PRINCIPLES-GUIDED LEARNING FOR PHYSICALLY PLAUSIBLE VIDEO GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Text-to-video (T2V) models have achieved remarkable visual fidelity, yet their generated videos still frequently violate physical laws. Preference-based post-training provides a promising way to improve physical plausibility, but existing approaches often represent physics supervision using a scalar signal, which collapses the principle-wise structure of physical behavior and provides limited information about how preferred and rejected videos differ physically. We introduce PhysGuide, a physics principles-guided preference optimization algorithm for physically plausible video generation. We construct a set of physics principle vectors from real-world demonstrations and project preferred and rejected videos into this representation space. PhysGuide uses their principle-wise differences to independently modulate the preferred and rejected policy rewards, producing physics-guided rewards along individual physical dimensions. We further introduce a physics-violation divergence penalty that captures the overall magnitude of physical disagreement between the pair, so that more severe violations produce a stronger corrective signal during optimization. We evaluate PhysGuide on PhyGenBench and VideoPhy and observe consistent improvements over the baselines, while preserving general video quality on VBench. Ablations further show that the principle-specific physics-guided rewards and physics-violation divergence penalty contribute complementary improvements to physical plausibility.

1 INTRODUCTION

Text-to-video (T2V) generation has advanced rapidly, with recent models (Yang et al., 2024; Wan et al., 2025; Kong et al., 2024) producing clips of striking visual fidelity. However, generated videos still lack physical realism and often violate basic physical laws: objects may pass through one another, fluids may defy gravity, or models may generate implausible contact dynamics. This shortcoming limits the utility of T2V models for downstream applications such as robotics, simulation, and virtual reality, where physical plausibility is a prerequisite rather than a stylistic preference.

Post-training has emerged as an effective approach for improving T2V quality, with Direct Preference Optimization (DPO) (Rafailov et al., 2023) and its diffusion variants (Liu et al., 2025b; Wallace et al., 2024), GRPO (Xue et al., 2025b; Zhang et al., 2026), and RLVR (Le et al., 2025) showing substantial gains in visual quality. Recent works (Chen et al., 2025; Cai et al., 2025; Zhang et al., 2026; Le et al., 2025; Ji et al., 2025) have explored adapting these methods to improve physical realism in video generation. Despite these efforts, gains on physics-centered benchmarks remain limited, and the gap between visual realism and physical correctness persists. We argue that this gap stems in part from how post-training methods represent and use physics information during optimization.

Why is physics-aware post-training difficult? Existing methods typically represent physics supervision using a scalar score (Cai et al., 2025; Wang et al., 2025b; Liu et al., 2025b). However, a scalar score reduces rich, multi-dimensional physics information to a single value and discards the principle-wise structure underlying the physical behavior. Different failures such as incorrect gravity, object penetration, wrong collision dynamics, or unrealistic fluid motion can therefore be represented in the same scalar form, even though they arise from different physical principles. This becomes particularly limiting when a video exhibits multiple physical phenomena or failures simultaneously. For example, two videos may receive similar physics scores while differing for orthogonal reasons, such as incorrect collision dynamics in one and inconsistent gravity in another.

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.

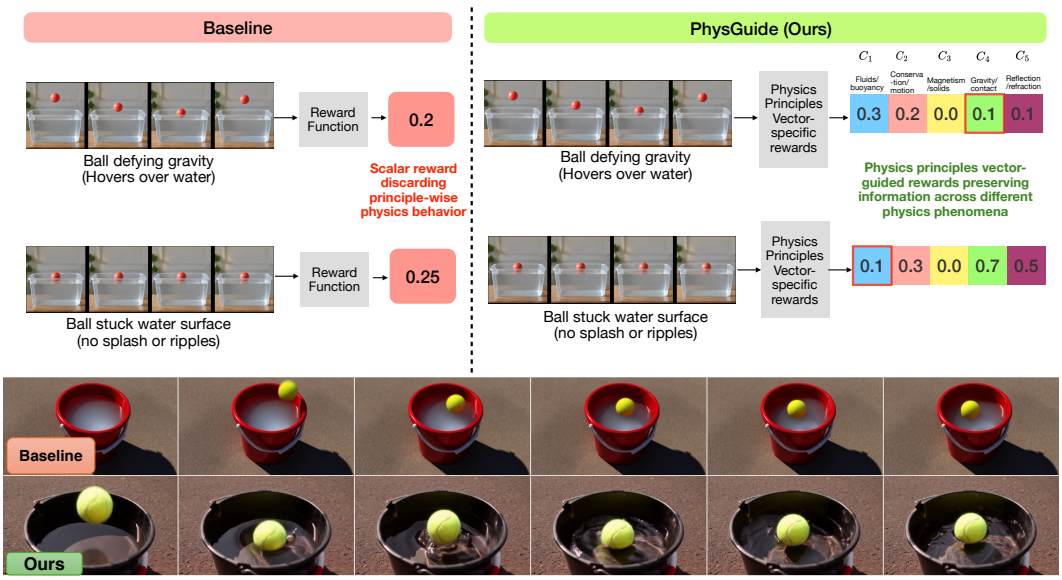

Figure 1: **Motivation and overview of PhysGuide.** Existing physics-aware post-training methods often compress physical behavior into a single scalar signal, obscuring which physical principles distinguish preferred and rejected videos. In contrast, PhysGuide represents videos in a physics-principle space and uses principle-wise differences to construct physics-guided policy rewards, while a physics-violation divergence penalty captures the overall magnitude of physical disagreement. This structured supervision helps the model learn more physically plausible interactions, such as realistic water displacement and ripples when the ball contacts the surface.

Beyond distinguishing the physical principles involved, post-training should also account for the severity of their disagreement. Physical violations can range from small deviations from expected dynamics to failures that fundamentally contradict the underlying interaction. For example, a ball that slightly penetrates a surface before rebounding represents a much milder violation than one that passes completely through the surface. A single scalar score may not preserve this structure or remain proportional to the underlying physical discrepancy. Prior works (Wang et al., 2025b; Cai et al., 2025; Liu et al., 2025b) address this primarily through sample weighting, using scalar physics scores to determine how strongly each sample contributes to optimization. However, sample weighting only changes the overall strength of a training example; it does not determine how the preference update should act along different physical principles or how it should modulate the preferred and rejected policy rewards. Thus, scalar supervision provides a coarse measure of overall physical plausibility but offers limited structure for principle-guided preference optimization.

To address these limitations, we introduce **PhysGuide**, a physics-principles-guided post-training algorithm for physically plausible text-to-video generation. Given a collection of real-world videos (Motamed et al., 2026) covering diverse physical phenomena, we identify the physics principles present in each video, group related behaviors, and construct a set of physics principle vectors that define a principle-wise representation space. We project each winner–loser video pair into this space to obtain physics-grounded representations of their physical behavior before preference-based post-training (Rafailov et al., 2023). In parallel, we compute the implicit policy rewards of the preferred and rejected videos under the trainable and reference policies. PhysGuide uses the principle-wise differences between the two video representations to independently modulate these policy rewards, producing physics-guided rewards for each principle. We further introduce a physics-violation divergence penalty based on the overall distance between the winner and loser representations, such that larger physical disagreements produce a stronger corrective signal during optimization.

The physics principles vector representation provides many advantages for physics-guided preference learning. A video can simultaneously involve gravity, collision, contact, deformation, fluid behavior, conservation, and other physical phenomena. Unlike a scalar score that compresses the physical behavior of a video into a single value as shown in Figure 1, our representation preserves its association with multiple physics dimensions. This allows PhysGuide to retain which principles distinguish the preferred and rejected videos and to use each dimension separately during preference optimization.

Consequently, differences along one physical principle can modulate the policy rewards differently from differences along another, rather than collapsing all physical information into a single training signal. At the same time, the physics-violation divergence penalty captures the overall magnitude of disagreement between the pair, so more severe physical disagreements produce a stronger corrective signal while milder deviations contribute proportionally less. In this way, PhysGuide captures both the principle-wise structure and the overall magnitude of physical disagreement between preference samples.

We evaluate PhysGuide on the physics-centered benchmarks PhyGenBench (Meng et al., 2024) and VideoPhy (Bansal et al., 2024), together with general video quality on VBench (Huang et al., 2024), and show improved performance over the baselines.

2 Related Works

Physics-aware Video Generation Video generation models have made significant progress in recent years, however a gap between visual and physics realism remains (Kang et al., 2024; Motamed et al., 2025; Meng et al., 2024). Benchmark evaluations (Bansal et al., 2024; Li et al., 2025b; Meng et al., 2024; Motamed et al., 2025; Wang et al., 2025a; Xiang et al., 2025) consistently show that current models still violate physics dynamics despite generating visually convincing videos. Towards improving the physical plausibility of generation, approaches targeting image-to-video generation have focused on identifying objects and modeling their motion from images/videos (Le Cleac'h et al., 2023; Montanaro et al., 2024; Yang et al., 2025; Liu et al., 2024; Li et al., 2025a). However, these methods are often specialized to particular objects, interaction types, or physical settings, which can limit their applicability to broader text-to-video generation. More recent works have targeted broader text-to-video generation by using VLM/LLM such as PhyT2V (Xue et al., 2025a), PhyRPR (Zhao et al., 2026), MM-CoT (Liu et al., 2025c), PhyMagic (Meng et al., 2025), or learning visual tokens (Lin et al., 2025). Other methods use parametric control and physics-informed Neural ODE (Yuan et al., 2025) or reason about implausible alternatives during denoising (Wang et al., 2026). In contrast, we construct physics principle vectors from real-world demonstrations and use them to represent generated videos along multiple physical dimensions during preference post-training.

Post-training for Video Generation Reinforcement learning (RL) post-training methods and reward design has been explored for physics plausible video generation by works such as PhysRVG (Zhang et al., 2026), NewtonRewards (Le et al., 2025), PhysMaster (Ji et al., 2025), PhysCorr (Wang et al., 2025b) PhysHPO (Chen et al., 2025), and PhysGDPO (Cai et al., 2025). Direct Preference Optimization (Rafailov et al., 2023) and its adaptation to diffusion models in Diffusion-DPO (Wallace et al., 2024) made preference-based post-training practical for visual generation, and subsequent work extended this idea to videos such as VideoDPO (Liu et al., 2025b). Other works (Jiang et al., 2025; Liu et al., 2025a) investigate using DPO to improve alignment with human preferences (Cheng et al., 2025), as online DPO (Zhang et al., 2024), for self-retrospective learning (Ahn et al., 2025), or group-based preference for DPO (Cai et al., 2025) using VLM-based rewards. Other works (Wu et al., 2024; Hong et al., 2026; Rho, 2025) have also explored margin-based formulation of DPO, however it is not straightforward to apply a general-purpose margin-based DPO formulation for learning physics. Closest to our work are PhysHPO (Chen et al., 2025) and PhyGDPO (Cai et al., 2025). While these methods incorporate physics-aware signals into preference learning, they reduce the physics information used during optimization to a scalar pairwise signal, providing limited structure for distinguishing which physical principles separate the preferred and rejected videos. Our work instead preserves principle-wise physics information through physics principle vectors, uses it to separately modulate the preferred and rejected policy rewards along individual physical dimensions, and captures the overall magnitude of physical disagreement through a physics-violation divergence penalty.

3 Method

We introduce PhysGuide, a physics principles-guided preference optimization framework for improving the physical plausibility of text-to-video generation. We first construct a physics-principle representation space from real-world videos using a set of physics principle vectors (Sec. 3.1). We then use these representations to modulate the preferred and rejected policy rewards along individual physics dimensions and combine them with a physics-violation divergence penalty in our final preference objective (Sec. 3.2). Figure 2 provides an overview of the approach.

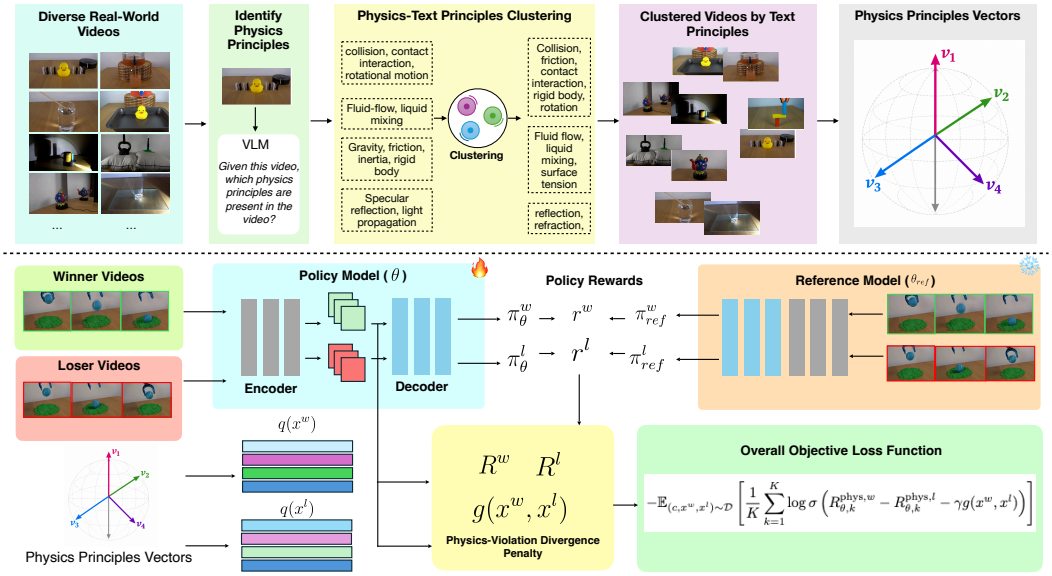

Figure 2: **(Row 1) Physics principle vector construction.** Given a collection of diverse real-world videos, a VLM identifies the dominant physical principles in each clip (e.g., collision, rigid motion, momentum conservation, mirror reflection, and fluid viscosity). We group videos according to their textual physics descriptions into K broad categories and aggregate the visual features within each group to initialize the corresponding physics principle vectors. **(Row 2) Overview of PhysGuide.** We project the preferred and rejected videos into the physics-principle space to obtain their principle-wise representations $\mathbf{q}^w$ and $\mathbf{q}^l$. In parallel, the trainable policy and frozen reference policy provide the implicit policy rewards for both videos. PhysGuide uses the principle-wise differences between $\mathbf{q}^w$ and $\mathbf{q}^l$ to construct physics-guided rewards for the preferred and rejected videos, while their squared ℓ_2 distance defines the physics-violation divergence penalty $g(x^w, x^l)$. We combine the physics-guided rewards and divergence penalty in the final preference objective to optimize the trainable policy.

3.1 Physics Principles Vectors

We represent physics principles using a set of vectors $\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_K\} \subset \mathbb{R}^d$, where each vector $\mathbf{p}_k$ represents a broad physical phenomenon. Together, these vectors define a physics-principle space in which a video can be represented according to its association with multiple underlying physical behaviors (Figure 2, Row 1).

Given a collection of real-world videos $\mathcal{V} = \{v_1, \ldots, v_N\}$ spanning diverse physical phenomena such as gravity, collision, fluid dynamics, and optical interactions, we first identify the physical principles exhibited by each video. For each video v_i, a vision-language model (VLM) produces a structured description s_i consisting of (1) a list of physical principles present in the video and (2) a summary describing the physical behavior over the course of the clip. Each principle is associated with a name, a category drawn from a closed taxonomy including mechanics, optics, fluids, thermal phenomena, electromagnetism, and material interactions, and a textual description of how the principle appears in the video.

We encode the textual description s_i using a sentence encoder f_{text} and ℓ_2-normalize the resulting embedding to obtain a physics-text embedding $\mathbf{e}_i$. We then partition these embeddings into K groups using K-means. We perform clustering entirely in the textual space of physics semantics, such that videos are grouped according to the physical phenomena they exhibit rather than their appearance, background, or object identity:

$$s_i = \text{VLM}(v_i), \qquad \mathbf{e}_i = \frac{f_{\text{text}}(s_i)}{\|f_{\text{text}}(s_i)\|_2}, \qquad \{\mathcal{C}_1, \ldots, \mathcal{C}_K\} = \text{KMeans}\left(\{\mathbf{e}_i\}_{i=1}^N\right). \tag{1}$$

For each cluster $\mathcal{C}_k$, we pass its corresponding videos through the video encoder f_ϕ to obtain visual features $\mathbf{z}_i \in \mathbb{R}^d$. We finally initialize the physics principle vector $\mathbf{p}_k$ by averaging the visual features of the videos assigned to that cluster:

$$\mathbf{z}_i = f_\phi(v_i), \qquad \mathbf{p}_k = \frac{1}{|\mathcal{C}_k|} \sum_{i \in \mathcal{C}_k} \mathbf{z}_i. \tag{2}$$

This initialization grounds each physics principle vector in visual demonstrations that share a common physical phenomenon.

Physics-principle projection of a video. Given an arbitrary video x, we encode it using the same video encoder, $\mathbf{z} = f_\phi(x)$, and compare its representation with each physics principle vector. We convert the resulting similarities into a normalized soft assignment over the K principles:

$$\mathbf{q}(x) = \mathrm{softmax}\left(\frac{[\mathrm{sim}(\mathbf{z}, \mathbf{p}_1), \ldots, \mathrm{sim}(\mathbf{z}, \mathbf{p}_K)]}{\tau}\right), \tag{3}$$

where $\mathrm{sim}(\cdot, \cdot)$ denotes cosine similarity and τ is a temperature parameter. The resulting $\mathbf{q}(x) = [q_1(x), \ldots, q_K(x)] \in \Delta^{K-1}$ provides a principle-wise projection of the video in the physics principle vector space, where $q_k(x)$ measures its relative association with physics principle vector $\mathbf{p}_k$.

The physics principle vectors and video representations share the embedding space of f_ϕ. During post-training, the representation space of the corresponding video encoder evolves together with the generator. To keep the physics principle vectors aligned with this evolving feature space, we maintain an exponential-moving-average (EMA) copy of the encoder and use it to update the physics principle vectors throughout training. The EMA update allows the vectors to track the evolving video features while suppressing short-term optimization noise, providing a stable principle vectors-based encoding for the training videos.

3.2 Physics Principles-Guided Preference Optimization

Our goal is to post-train a text-to-video generation policy using preference supervision that is explicitly grounded in the physics principle representations introduced in Sec. 3.1. We consider a preference dataset $\mathcal{D} = \left\{(c_i, x_i^w, x_i^l)\right\}_{i=1}^N$, where each dataset sample pairs a text prompt c_i with a physically preferred video x_i^w and a rejected video x_i^l exhibiting physically implausible behavior.

Physics-principle representations of preference pairs. Given a winner–loser pair (x^w, x^l), we project both videos into the physics-principle space defined in Sec. 3.1 to obtain $\mathbf{q}^w = \mathbf{q}(x^w)$, $\mathbf{q}^l = \mathbf{q}(x^l)$, where $\mathbf{q}^w = [q_1^w, \ldots, q_K^w]$ and $\mathbf{q}^l = [q_1^l, \ldots, q_K^l]$ are normalized distributions over the K physics principle vectors. Rather than immediately reducing the difference between these representations to a single scalar, we retain their individual principle dimensions throughout preference optimization.

Policy rewards. In parallel, we evaluate the preferred and rejected videos under the trainable policy π_θ relative to the frozen reference policy π_{ref}. Following the implicit reward formulation of direct preference optimization, we define

$$r_\theta(x, c) = \beta \log \frac{\pi_\theta(x \mid c)}{\pi_{\mathrm{ref}}(x \mid c)}, \qquad r_\theta^w = r_\theta(x^w, c), \qquad r_\theta^l = r_\theta(x^l, c), \tag{4}$$

where β controls the strength of deviation from the reference policy. We compute these rewards from the full text-to-video policy, and they quantify how much more strongly the learned policy favors each video relative to the reference model.

Principle-specific policy rewards. The scalar policy rewards r_θ^w and r_θ^l do not distinguish which physical principles separate the preferred and rejected videos. We therefore use their physics-principle

representations to modulate the two rewards independently. For each principle k, we define the signed principle discrepancy

$$d_k = q_k^w - q_k^l. \tag{5}$$

We then construct the principle-specific physics-guided policy rewards as

$$R_{\theta,k}^{\text{phys},w} = (1 + \lambda[d_k]_+)\, r_\theta^w, \qquad R_{\theta,k}^{\text{phys},l} = (1 + \lambda[-d_k]_+)\, r_\theta^l. \tag{6}$$

where $[u]_+ = \max(u, 0)$ and $\lambda \geq 0$ controls the strength of the principle-specific modulation. When $q_k^w > q_k^l$, principle k increases the contribution of the preferred-video reward; when $q_k^l > q_k^w$, it instead increases the contribution of the rejected-video reward. When the two videos have similar association with principle k, both modulation factors remain close to one.

This construction allows each physics dimension to independently determine how the optimization allocates the preference update between the preferred and rejected videos, while the underlying implicit rewards r_θ^w and r_θ^l remain those of the full text-to-video policy. Setting $\lambda = 0$ gives $R_{\theta,k}^{\text{phys},w} = r_\theta^w$ and $R_{\theta,k}^{\text{phys},l} = r_\theta^l$ for every principle.

Physics-violation divergence penalty. In addition to the principle-specific reward modulation, we retain a global measure of the physics discrepancy between the preferred and rejected videos. Given their physics-principle representations $\mathbf{q}^w$ and $\mathbf{q}^l$, we define the physics-violation divergence penalty as the squared ℓ_2 distance between the two distributions:

$$g(x^w, x^l) = \left\| \mathbf{q}^w - \mathbf{q}^l \right\|_2^2. \tag{7}$$

The divergence is small when the preferred and rejected videos exhibit similar physical behavior and increases as they differ across the physics-principle dimensions. We use this quantity as a physics-violation divergence penalty, so that more severe physical disagreements produce a stronger corrective signal during preference optimization, while milder deviations contribute proportionally less.

Physics principles-guided preference objective. Combining the principle-specific policy rewards with the physics-violation divergence penalty gives our final preference objective (Figure 2, Row 2),

$$\mathcal{L}_{\text{PhysGuide}} = -\mathbb{E}_{(c,x^w,x^l)\sim\mathcal{D}} \left[\frac{1}{K} \sum_{k=1}^{K} \log \sigma \left(R_{\theta,k}^{\text{phys},w} - R_{\theta,k}^{\text{phys},l} - \gamma g(x^w, x^l) \right) \right]. \tag{8}$$

where $\gamma \geq 0$ controls the strength of the physics-violation divergence penalty. The principle-specific physics-guided rewards and the divergence penalty play complementary roles: $R_{\theta,k}^{\text{phys},w}$ and $R_{\theta,k}^{\text{phys},l}$ determine how the preference signal is allocated between the preferred and rejected videos along each physics principle, while $g(x^w, x^l)$ controls the strength of the corrective signal for the pair.

Setting $\lambda = 0$ removes the principle-specific modulation and recovers the original scalar policy rewards for every principle. Setting $\gamma = 0$ removes the physics-violation divergence penalty. Together, the two terms allow the physics-principle representation to influence both the allocation and the magnitude of the preference update while leaving the underlying text-to-video policy architecture unchanged.

4 EXPERIMENTS

Baselines and ablations We compare PhysGuide with the baselines, PhyT2V (Xue et al., 2025a), PhysHPO (Chen et al., 2025), and PhyGDPO (Cai et al., 2025). In the ablation study, we analyze the contributions of the physics principle vectors, principle-specific physics-guided rewards, physics-violation divergence penalty, and EMA updates to the final performance of the model.

Method	PhyGenBench					VideoPhy			
	Mechanics	Optics	Thermal	Material	Overall	SS	SF	FF	Overall
CogVideoX-2B (Yang et al., 2024)	38.0	43.0	34.0	39.0	39.0	12.7	21.9	25.4	18.6
CogVideoX-5B (Yang et al., 2024)	39.0	55.0	40.0	42.0	45.0	24.4	53.1	43.6	39.6
HunyuanVideo	37.5	58.0	36.7	45.0	45.6	19.7	43.2	42.9	33.4
PhyT2V (Xue et al., 2025a)	40.0	58.0	33.3	47.5	46.3	21.1	45.9	50.0	36.3
PhysHPO (Chen et al., 2025)	47.5	62.0	43.3	60.0	54.4	26.8	50.7	55.4	41.6
PhysGDPO (Chen et al., 2025)	55.0	60.0	58.0	47.0	55.0	-	-	-	-
PhysGuide (Ours)	55.5	72.1	60.6	61.5	63.1	32.2	55.1	59.2	45.3

Table 1: **Main results on PhyGenBench and VideoPhy.** We compare PhysGuide with base T2V models and physics-aware video generation methods. For VideoPhy, SS, SF, and FF denote solid–solid, solid–fluid, and fluid–fluid interactions, respectively. Higher scores indicate better physical plausibility on both benchmarks. PhysGuide achieves the best overall performance on both benchmarks, with consistent gains across physics categories and interaction types.

Preference Post-Training	Physics-Guided Rewards	Divergence Penalty	PhyGenBench					VideoPhy			
			Mech	Opt	Therm	Mat	Over	SS	SF	FF	Over
			37.5	58.0	36.7	45.0	45.6	19.7	43.2	42.9	33.4
✓			46.1	61.0	42.1	58.7	53.1	23.2	49.3	54.2	40.7
✓	✓		52.2	67.4	54.2	60.1	59.3	28.1	52.4	56.7	43.6
✓	✓	✓	55.5	72.1	60.6	61.5	63.1	32.2	55.1	59.2	45.3

Table 2: **Ablation study of PhysGuide.** Each row progressively adds components to the base HunyuanVideo model. We evaluate preference post-training, the principle-specific physics-guided policy rewards (Physics-Guided Rewards), and the physics-violation divergence penalty (Divergence Penalty). Physics-guided rewards provide substantial gains over standard preference post-training, while the divergence penalty further improves physical plausibility on both benchmarks.

Evaluation benchmarks. We evaluate PhysGuide on two physics-centered benchmarks. **PhyGenBench** Meng et al. (2024) consists of 160 prompts across 27 physical laws and four domains (mechanics, optics, thermal, and material properties) to evaluate "physical commonsense" in video generation. It measures physical plausibility across the four domains and reports an aggregate overall score. **VideoPhy** Bansal et al. (2024) utilizes a human-verified dataset of 688 captions covering object interaction states such as solid-solid, solid-fluid, and fluid-fluid. It evaluates physical consistency of solid–solid, solid–fluid, and fluid–fluid interactions, with an aggregate overall score. We further evaluate the overall visual and semantic quality of the generated videos on the **VBench** (Huang et al., 2024) benchmark. Higher is better on all benchmarks.

Implementation Details. We use HunyuanVideo as our text-to-video (T2V) generation model. We use the Physics-IQ dataset (Motamed et al., 2026) for finetuning and train PhysGuide with AdamW using a learning rate of 5×10^{-6} on one NVIDIA A100. We use $K = 8$ physics principle vectors, where each vector is initialized from a cluster of real-world physics demonstrations. We discuss the construction of the physics principle vectors in the Appendix.

4.1 Main Results

Referring to Table 1, PhysGuide outperforms the baselines on the PhyGenBench, VideoPhy, and VBench benchmarks. We attribute these improvements to the physics principle representation and its integration into preference optimization through principle-specific physics-guided rewards and the physics-violation divergence penalty, which together provide a more structured learning signal for distinguishing physically plausible and implausible dynamics.

Improvement in physics plausibility. We compare PhysGuide comprehensively with the baselines, PhyT2V (Xue et al., 2025a), PhysHPO (Chen et al., 2025), and PhyGDPO (Cai et al., 2025) in Table 1, together with the qualitative results in Figure 4 and the human study in Figure 3. As shown in Table 1, PhysGuide consistently performs better than the physics-based baselines across the evaluated benchmarks. These results indicate that retaining principle-wise physics information and incorporating it into the preferred and rejected policy rewards, together with the physics-violation divergence penalty, improves physical plausibility compared with methods that rely on scalar physics supervision like PhysHPO. PhysGuide also preserves visual and semantic quality on VBench (Huang et al., 2024), as shown in Table 1.

Method	PhyGenBench					VideoPhy			
	Mechanics	Optics	Thermal	Material	Over.	Solid-Solid	Solid-Fluid	Fluid-Fluid	Over.
Ours ($K = 4$)	53.2	68.5	56.4	60.2	60.3	30.7	53.6	57.5	44.1
PhysicsIQ ($K = 5$)	53.4	68.9	57.3	60.2	60.7	30.9	53.6	57.6	44.2
Ours ($K = 5$)	54.8	70.7	59.0	61.0	62.1	31.1	53.8	57.9	44.4
Random ($K = 8$)	49.5	65.1	47.4	59.6	56.5	25.2	49.6	54.8	40.6
Ours ($K = 8$)	55.5	72.1	60.6	61.5	63.1	32.2	55.1	59.2	45.3
Ours ($K = 16$)	54.1	71.5	60.2	60.9	62.3	31.5	54.3	58.1	44.7

Table 3: **Ablation of the physics principle vectors.** We study the effect of the number of physics principle vectors K and the strategy used to group real-world videos for their initialization. Physics-based grouping consistently outperforms random grouping, and $K = 8$ achieves the strongest overall performance on both benchmarks.

Method	VBench		
	Total	Quality	Semantics
CogVideoX-2B	81.6	82.5	77.8
CogVideoX-5B	81.9	83.1	77.3
HunyuanVideo	81.4	83.1	74.9
PhyT2V (Xue et al., 2025a)	80.5	82.0	74.5
PhysHPO (Chen et al., 2025)	81.9	83.5	75.7
PhysGuide (Ours)	84.3	85.9	77.9

Table 4: **General video quality on VBench.** PhysGuide achieves the highest Total and Quality scores while maintaining strong semantic consistency, showing that improved physical plausibility does not come at the expense of general video quality.

Figure 4 presents qualitative comparisons across the evaluated methods. For the tennis-ball example, PhysGuide generates the expected ripple pattern as the ball contacts the water surface, whereas the competing generations exhibit physically implausible behavior such as hovering or missing surface interactions. For the paintbrush example, PhysGuide produces yellow paint that follows the brush trajectory on the canvas, while the competing generations exhibit missing or inconsistent paint strokes. We also conduct a human preference study for physics plausibility using prompts from PhyGenBench and Impossible-Videos (Bai et al., 2025). Figure 3 shows that participants preferred PhysGuide in 96% of comparisons against HunyuanVideo and 78% of comparisons against HunyuanVideo + DPO on the PhyGenBench prompts. For the Impossible-Videos prompts, participants preferred PhysGuide over the base HunyuanVideo model in 75% of comparisons, further supporting the improvements in physical plausibility over the random 50% baseline.

4.2 Ablation Study

How does the number of physics vectors affect performance? Referring to Table 3, reducing K from 8 to 5 and 4 degrades performance because fewer vectors must represent a broader range of distinct physical behaviors, causing individual vectors to span multiple physics dimensions. Increasing K to 16 yields no significant gain over $K = 8$, indicating that $K = 8$ provides sufficient coverage of the physics dimensions represented in the preference data.

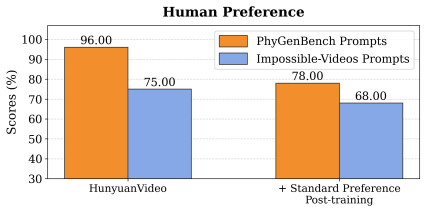

Figure 3: **Human preference study on PhyGenBench and Impossible-Videos prompts.** Participants prefer PhysGuide in 96% of comparisons against HunyuanVideo and 78% against preference post-training on PhyGenBench prompts, and in 75% of comparisons against HunyuanVideo on Impossible-Videos prompts.

Do different video grouping initialization strategies affect performance? We analyze random initialization by assigning videos randomly to $K = 8$ groups. As shown in Table 3, random grouping significantly degrades performance, demonstrating the importance of grouping videos according to meaningful physical behavior when constructing the physics principle vectors. These results show that clustering videos according to their textual descriptions of physical phenomena provides a more effective initialization of the physics principle vectors.

Our grouping mechanism vs. the dataset's predefined grouping. We compare our grouping strategy with the predefined Physics-IQ grouping at $K = 5$. For a controlled comparison, Table 3 compares both strategies using the same number of groups. Our clustering strategy performs better than the predefined Physics-IQ grouping, showing that the grouping mechanism itself contributes to the performance gain. Together with the results across different values of K, this suggests that the improvement arises from both the physics-based grouping strategy and the choice of the number of physics principle vectors.

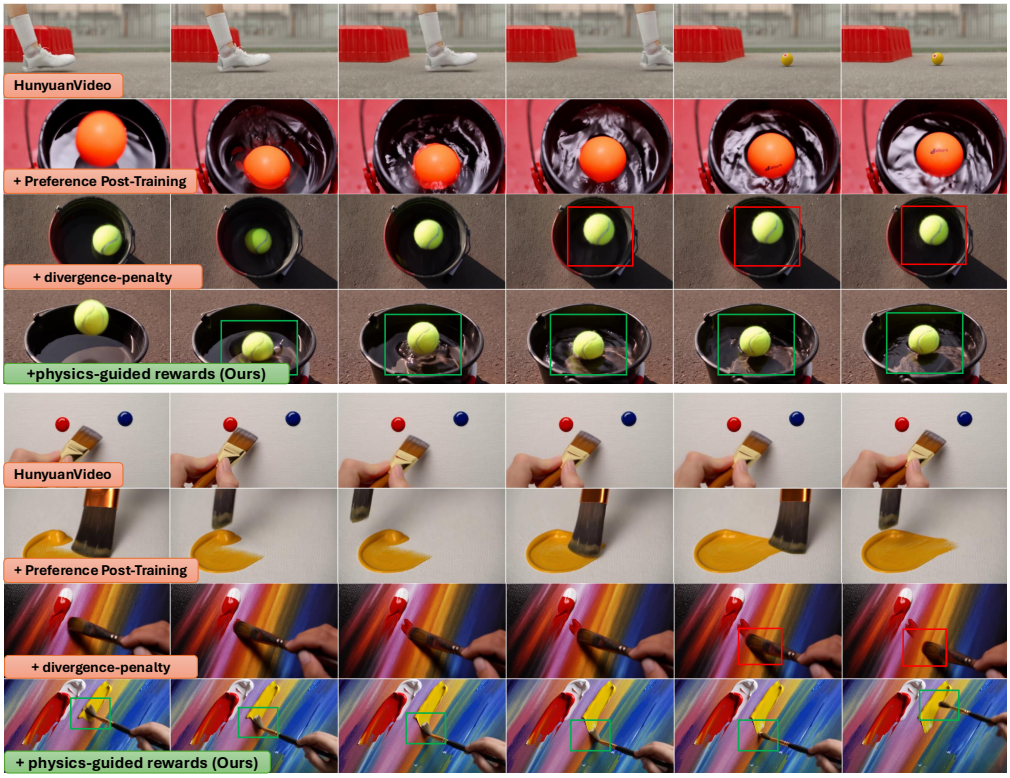

Figure 4: Qualitative results of PhysGuide and the ablation study. Text prompts: **(Top)** `A tennis ball is gently placed on the surface of a bucket filled with water.` **(Bottom)** `Bristles of paintbrush sweeping across canvas.`

Physics-guided rewards and divergence penalty improve physics learning. Table 2 progressively incorporates the components of PhysGuide into the base HunyuanVideo model. Compared to standard preference post-training, introducing the principle-specific physics-guided policy rewards significantly improves the overall PhyGenBench score from 53.1 to 59.3 and the overall VideoPhy score from 40.7 to 43.6. This result shows that using the physics principle representations to separately modulate the preferred and rejected policy rewards provides a stronger physics-aware preference signal than standard preference post-training alone. Adding the physics-violation divergence penalty produces a further improvement, raising the overall scores to 63.1 on PhyGenBench and 45.3 on VideoPhy. This gain indicates that, in addition to preserving principle-wise differences, accounting for the overall magnitude of physical disagreement between the preferred and rejected videos provides a stronger corrective signal during preference optimization. The improvement corroborates the qualitative results (Figure 4) where the red brush moves along with the brush trajectory unlike standard preference optimization where the brush simply hovers over the canvas.

5 Conclusion

We present PhysGuide, a physics principles-guided preference post-training method for physically plausible text-to-video generation. Motivated by the observation that scalar physics supervision compresses rich, multi-dimensional physical behavior into a single value, we introduce physics principle vectors constructed from clustered real-world demonstrations. PhysGuide uses these representations to modulate the preferred and rejected policy rewards along different physics dimensions, producing principle-specific physics-guided rewards, while a physics-violation divergence penalty captures the overall magnitude of physical disagreement between the pair. On PhyGenBench and VideoPhy, PhysGuide outperforms the baselines across the evaluated categories, and our ablations show that the physics-guided rewards and divergence penalty provide complementary gains. A limitation of the current approach is that the physics principle vectors derive from the physical phenomena represented in the initialization corpus, which may limit coverage of behaviors outside this distribution. Future work can extend the physics principle representation to hierarchical or dynamically expandable structures that capture a broader range of physical phenomena.

REFERENCES

Daechul Ahn, Yura Choi, San Kim, Youngjae Yu, Dongyeop Kang, and Jonghyun Choi. Isr-dpo: Aligning large multimodal models for videos by iterative self-retrospective dpo. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 1728–1736, 2025.

Zechen Bai, Hai Ci, and Mike Zheng Shou. Impossible videos. *arXiv preprint arXiv:2503.14378*, 2025.

Hritik Bansal, Zongyu Lin, Tianyi Xie, Zeshun Zong, Michal Yarom, Yonatan Bitton, Chenfanfu Jiang, Yizhou Sun, Kai-Wei Chang, and Aditya Grover. Videophy: Evaluating physical commonsense for video generation. *arXiv preprint arXiv:2406.03520*, 2024.

Yuanhao Cai, Kunpeng Li, Menglin Jia, Jialiang Wang, Junzhe Sun, Feng Liang, Weifeng Chen, Felix Juefei-Xu, Chu Wang, Ali Thabet, Xiaoliang Dai, Xuan Ju, Alan L. Yuille, and Ji Hou. Phygdpo: Physics-aware groupwise direct preference optimization for physically consistent text-to-video generation. *ArXiv*, abs/2512.24551, 2025. URL https://api.semanticscholar.org/CorpusID:284350537.

Harold Haodong Chen, Haojian Huang, Qifeng Chen, Harry Yang, and Ser-Nam Lim. Hierarchical fine-grained preference optimization for physically plausible video generation. *arXiv preprint arXiv:2508.10858*, 2025.

Haoran Cheng, Qide Dong, Liang Peng, Zhizhou Sha, Weiguo Feng, Jinghui Xie, Zhao Song, Shilei Wen, Xiaofei He, and Boxi Wu. Discriminator-free direct preference optimization for video diffusion. *arXiv preprint arXiv:2504.08542*, 2025.

Jiwoo Hong, Sayak Paul, Noah Lee, Kashif Rasul, James Thorne, and Jongheon Jeong. Margin-aware preference optimization for aligning diffusion models without reference. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 4744–4752, 2026.

Ziqi Huang, Yinan He, Jiashuo Yu, Fan Zhang, Chenyang Si, Yuming Jiang, Yuanhan Zhang, Tianxing Wu, Qingyang Jin, Nattapol Chanpaisit, et al. Vbench: Comprehensive benchmark suite for video generative models. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21807–21818. IEEE, 2024.

Sihui Ji, Xi Chen, Xin Tao, Pengfei Wan, and Hengshuang Zhao. Physmaster: Mastering physical representation for video generation via reinforcement learning. *ArXiv*, abs/2510.13809, 2025. URL https://api.semanticscholar.org/CorpusID:282102265.

Lifan Jiang, Boxi Wu, Jiahui Zhang, Xiaotong Guan, and Shuang Chen. Huvidpo: Enhancing video generation through direct preference optimization for human-centric alignment. *arXiv preprint arXiv:2502.01690*, 2025.

Bingyi Kang, Yang Yue, Rui Lu, Zhijie Lin, Yang Zhao, Kaixin Wang, Gao Huang, and Jiashi Feng. How far is video generation from world model: A physical law perspective. *arXiv preprint arXiv:2411.02385*, 2024.

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.

Minh-Quan Le, Yuanzhi Zhu, Vicky Kalogeiton, and Dimitris Samaras. What about gravity in video generation? post-training newton's laws with verifiable rewards. *ArXiv*, abs/2512.00425, 2025. URL https://api.semanticscholar.org/CorpusID:283449158.

Simon Le Cleac'h, Hong-Xing Yu, Michelle Guo, Taylor Howell, Ruohan Gao, Jiajun Wu, Zachary Manchester, and Mac Schwager. Differentiable physics simulation of dynamics-augmented neural objects. *IEEE Robotics and Automation Letters*, 8(5):2780–2787, 2023.

Chenyu Li, Oscar Michel, Xichen Pan, Sainan Liu, Mike Roberts, and Saining Xie. Pisa experiments: Exploring physics post-training for video diffusion models by watching stuff drop. *arXiv preprint arXiv:2503.09595*, 2025a.

Dacheng Li, Yunhao Fang, Yukang Chen, Shuo Yang, Shiyi Cao, Justin Wong, Michael Luo, Xiaolong Wang, Hongxu Yin, Joseph E Gonzalez, et al. Worldmodelbench: Judging video generation models as world models. *arXiv preprint arXiv:2502.20694*, 2025b.

Wang Lin, Liyu Jia, Wentao Hu, Kaihang Pan, Zhongqi Yue, Wei Zhao, Jingyuan Chen, Fei Wu, and Hanwang Zhang. Reasoning physical video generation with diffusion timestep tokens via reinforcement learning. *ArXiv*, abs/2504.15932, 2025. URL https://api.semanticscholar.org/CorpusID:277993895.

Jie Liu, Gongye Liu, Jiajun Liang, Ziyang Yuan, Xiaokun Liu, Mingwu Zheng, Xiele Wu, Qiulin Wang, Menghan Xia, Xintao Wang, et al. Improving video generation with human feedback. *arXiv preprint arXiv:2501.13918*, 2025a.

Runtao Liu, Haoyu Wu, Ziqiang Zheng, Chen Wei, Yingqing He, Renjie Pi, and Qifeng Chen. Videodpo: Omni-preference alignment for video diffusion generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 8009–8019, 2025b.

Shaowei Liu, Zhongzheng Ren, Saurabh Gupta, and Shenlong Wang. Physgen: Rigid-body physics-grounded image-to-video generation. In *European Conference on Computer Vision*, pp. 360–378. Springer, 2024.

Yang Liu, Xilin Zhao, Peisong Wen, Siran Dai, and Qingming Huang. Bootstrapping physics-grounded video generation through vlm-guided iterative self-refinement. *ArXiv*, abs/2511.20280, 2025c. URL https://api.semanticscholar.org/CorpusID:283251139.

Fanqing Meng, Jiaqi Liao, Xinyu Tan, Wenqi Shao, Quanfeng Lu, Kaipeng Zhang, Yu Cheng, Dianqi Li, Yu Qiao, and Ping Luo. Towards world simulator: Crafting physical commonsense-based benchmark for video generation. *arXiv preprint arXiv:2410.05363*, 2024.

Siwei Meng, Yawei Luo, and Ping Liu. Phymagic: Physical motion-aware generative inference with confidence-guided llm. 2025. URL https://api.semanticscholar.org/CorpusID:278788960.

Antonio Montanaro, Luca Savant Aira, Emanuele Aiello, Diego Valsesia, and Enrico Magli. Motioncraft: Physics-based zero-shot video generation. *Advances in Neural Information Processing Systems*, 37:123155–123181, 2024.

Saman Motamed, Laura Culp, Kevin Swersky, Priyank Jaini, and Robert Geirhos. Do generative video models learn physical principles from watching videos? *arXiv e-prints*, pp. arXiv–2501, 2025.

Saman Motamed, Laura Culp, Kevin Swersky, Priyank Jaini, and Robert Geirhos. Do generative video models understand physical principles? In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 948–958, 2026.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741, 2023.

Hyung Gyu Rho. Margin adaptive dpo: Leveraging reward model for granular control in preference optimization. *arXiv preprint arXiv:2510.05342*, 2025.

Bram Wallace, Meihua Dang, Rafael Rafailov, Linqi Zhou, Aaron Lou, Senthil Purushwalkam, Stefano Ermon, Caiming Xiong, Shafiq Joty, and Nikhil Naik. Diffusion model alignment using direct preference optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8228–8238, 2024.

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

Cong Wang, Hanxin Zhu, Xiao Tang, Jiayi Luo, Xin Jin, Long Chen, Fei-Yue Wang, and Zhibo Chen. Physvideo: Physically plausible video generation with cross-view geometry guidance. *arXiv preprint arXiv:2603.18639*, 2026.

Jing Wang, Ao Ma, Ke Cao, Jun Zheng, Zhanjie Zhang, Jiasong Feng, Shanyuan Liu, Yuhang Ma, Bo Cheng, Dawei Leng, et al. Wisa: World simulator assistant for physics-aware text-to-video generation. *arXiv preprint arXiv:2503.08153*, 2025a.

Peiyao Wang, Weining Wang, and Qi Li. Physcorr: Dual-reward dpo for physics-constrained text-to-video generation with automated preference selection. *arXiv preprint arXiv:2511.03997*, 2025b.

Junkang Wu, Xue Wang, Zhengyi Yang, Jiancan Wu, Jinyang Gao, Bolin Ding, Xiang Wang, and Xiangnan He. Alphadpo: Adaptive reward margin for direct preference optimization. *arXiv preprint arXiv:2410.10148*, 2024.

Kun Xiang, Heng Li, Terry Jingchen Zhang, Yinya Huang, Zirong Liu, Peixin Qu, Jixi He, Jiaqi Chen, Yu-Jie Yuan, Jianhua Han, et al. Seephys: Does seeing help thinking?–benchmarking vision-based physics reasoning. *arXiv preprint arXiv:2505.19099*, 2025.

Qiyao Xue, Xiangyu Yin, Boyuan Yang, and Wei Gao. Phyt2v: Llm-guided iterative self-refinement for physics-grounded text-to-video generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 18826–18836, 2025a.

Zeyue Xue, Jie Wu, Yu Gao, Fangyuan Kong, Lingting Zhu, Mengzhao Chen, Zhiheng Liu, Wei Liu, Qiushan Guo, Weilin Huang, et al. Dancegrpo: Unleashing grpo on visual generation. *arXiv preprint arXiv:2505.07818*, 2025b.

Xindi Yang, Baolu Li, Yiming Zhang, Zhenfei Yin, Lei Bai, Liqian Ma, Zhiyong Wang, Jianfei Cai, Tien-Tsin Wong, Huchuan Lu, et al. Vlipp: Towards physically plausible video generation with vision and language informed physical prior. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 12360–12370, 2025.

Zhuoyi Yang, Jiayan Teng, Wendi Zheng, Ming Ding, Shiyu Huang, Jiazheng Xu, Yuanming Yang, Wenyi Hong, Xiaohan Zhang, Guanyu Feng, et al. Cogvideox: Text-to-video diffusion models with an expert transformer. *arXiv preprint arXiv:2408.06072*, 2024.

Yu Yuan, Xijun Wang, Tharindu Wickremasinghe, Zeeshan Nadir, Bole Ma, and Stanley H Chan. Newtongen: Physics-consistent and controllable text-to-video generation via neural newtonian dynamics. *arXiv preprint arXiv:2509.21309*, 2025.

Jiacheng Zhang, Jie Wu, Weifeng Chen, Yatai Ji, Xuefeng Xiao, Weilin Huang, and Kai Han. Onlinevpo: Align video diffusion model with online video-centric preference optimization. *arXiv preprint arXiv:2412.15159*, 2024.

Qiyuan Zhang, Biao Gong, Shuai Tan, Zheng Zhang, Yujun Shen, Xing Zhu, Yuyuan Li, Kelu Yao, Chunhua Shen, and Changqing Zou. Physrvg: Physics-aware unified reinforcement learning for video generative models. *arXiv preprint arXiv:2601.11087*, 2026.

Yibo Zhao, Hengjia Li, Xiaofei He, and Boxi Wu. Phyrpr: Training-free physics-constrained video generation. *ArXiv*, abs/2601.09255, 2026. URL `https://api.semanticscholar.org/CorpusID:284717757`.

A APPENDIX

A.1 PHYSICS PRINCIPLES VECTORS CURATION

Given N videos in the Physics-IQ dataset, we query the VLM (GPT-5.4) to query each video with the question - "Given this video, which physics principles are present in the video?". The VLM gives us the physics principles keywords in text format which are then clustered to form K codes. The videos with respect to them are clustered together and passed through the visual encoder. The embeddings of the clustered videos are mean-pooled together which represents each physics phenomenon of the vectors. The responses from the VLM are shared at the Appendix.

A.2 PREFERENCE DATASET CURATION

For a given caption c, we generate G losers using the video generation model. We keep $G = 3$ for our preference dataset generation. This gives us 200 winner videos and 600 loser videos. The G losers generated from the video generation model are manually checked and selected to cover diverse error types. For each prompt we first generate 10 candidate loser videos using different random seeds from the base video generation model. These candidates are then manually inspected to ensure that the selected losers cover diverse categories of physics failures (e.g., implausible motion, object disappearance, object hallucination, temporal motion inconsistencies, and other physical violations). The purpose of this manual selection is to ensure diversity of negative examples and exposes the model to a broader range of physics violations during training, instead of overfitting to a single failure type.

A.3 QUANTITATIVE RESULTS

Appearance Sensitivity Analysis. Our training pairs are specifically constructed to minimize appearance variation. For every winner video, the corresponding loser videos are generated from the same prompt and image of the winner video using I2V generation. This ensures that the winner and loser share the same scene configuration and that the primary difference between the pair is the physical dynamics rather than the visual appearance. As the physics-violation divergence penalty is calculated between the pairs which share the same appearance configuration, it reduces the possibility of conflating appearance differences with physical differences.

We evaluate the sensitivity of the physics-violation divergence penalty to the appearance changes. We selected six videos and created five appearance-only augmentations by applying a blue color tint, green color tint, red color tint uniformly to every frame of the video. This preserves the underlying physical dynamics while modifying the visual style. We then computed the Euclidean distance between (1) the first video and its augmented versions and (2) the first video and all the five other original videos.

Similarity Across Augmentations						
Reference Video	**Aug. 1**	**Aug. 2**	**Aug. 3**	**Aug. 4**	**Aug. 5**	**Average**
Video 1	0.04	0.07	0.03	0.04	0.05	0.046

Similarity Across Different Videos						
Reference Video	**Video 1**	**Video 2**	**Video 3**	**Video 4**	**Video 5**	**Average**
Video 1	0.53	0.47	0.89	0.35	0.68	0.584

Table 5: Similarity across augmentations and across different videos.

As shown in the Table 5, the physics-violation divergence penalty remains consistently small between a video and its appearance-only augmentations. This indicates that the physics-violation divergence penalty and its appearance-only augmentations will also have a small difference when computed against the corresponding winner sample. In contrast, the physics-violation divergence penalty becomes substantially larger between different videos where both the physical behavior and visual content differ. And so, their physics-violation divergence penalty with corresponding winner samples can also exhibit large differences. These results suggest that the proposed physics-violation divergence penalty is considerably more sensitive to changes in physical behavior across videos than to appearance variations.

VLM responses


```
{
 "principles": [
   {
     "name": "gravity",
     "category": "mechanics",
     "description": "A ball and a small block are released from above and accelerate
         downward onto the cushions due to gravity."
   },
   {
     "name": "free fall",
     "category": "mechanics",
     "description": "Before contact with the cushions, both objects undergo downward
         motion primarily under the influence of gravity."
   },
   {
     "name": "impact deformation",
     "category": "material",
     "description": "When the objects land, the soft cushions visibly compress and
         deform under the contact forces."
   },
   {
     "name": "energy dissipation",
     "category": "mechanics",
     "description": "The objects do not rebound much after landing, indicating that
         kinetic energy is absorbed by the soft cushion material and converted into
         deformation and heat."
   }
 ],
 "summary": "A ball and a small block are dropped simultaneously onto two soft
     cushions. They fall under gravity, then the cushions compress on impact and
     absorb most of the motion."
}
```


VLM responses


```
{
"principles": [
    {
     "name": "occlusion",
     "category": "optics",
     "description": "As the paper rotates in front of the ball, it blocks the line of
          sight so the ball disappears and reappears depending on whether it is
         hidden behind the paper."
    },
    {
     "name": "rotational motion",
     "category": "mechanics",
     "description": "The paper assembly turns about a vertical axis on the motorized
         base, changing its orientation relative to the camera over time."
    },
    {
     "name": "rigid body motion",
     "category": "mechanics",
     "description": "The folded paper moves as a single solid object, maintaining its
          shape while rotating around the stand."
    }
   ],
   "summary": "A folded sheet of paper rotates on a stand in front of a suspended ball.
        The ball is visible when not aligned behind the paper and becomes hidden when
       the rotating paper passes between it and the camera."
}
```

VLM responses


```
{
"principles": [
    {
     "name": "absorption and wicking",
     "category": "fluids",
     "description": "The paper towel rapidly takes up the blue liquid as its porous
          fibers wick water into the sheet."
    },
    {
     "name": "capillary action",
     "category": "fluids",
     "description": "Liquid is drawn through the tiny gaps between paper fibers,
          allowing the napkin to soak and spread the water through itself."
    },
    {
     "name": "gravity",
     "category": "mechanics",
     "description": "The napkin is released above the plate and falls downward into
          the liquid under gravity."
    },
    {
     "name": "porous material saturation",
     "category": "material",
     "description": "As the towel becomes saturated, its structure darkens and
          collapses flatter into the plate due to the added liquid mass and softened
          fibers."
}
```


A.4 AI ASSISTANCE

We used AI to fix grammatical errors and proof read the writing in the paper.