A Ethics Statement

This paper studies methods for preference optimization in diffusion models. Our contributions are mainly methodological. Potential positive impacts include improving controllability of generative models, which may benefit applications such as art creation and scientific visualization. However, misuse of the method could lead to harmful content generation or reinforce biases present in the training data. We rely only on publicly available datasets in our experiments, and we encourage future work to carefully consider fairness, safety, and data governance issues when deploying such models.

B Use of LLMs

We used large language models (LLMs) solely for grammar and spelling checking in the preparation of this paper. No LLMs were involved in generating research ideas, conducting experiments, analyzing results, or writing technical content. All scientific contributions and claims are entirely the work of the authors.

C Related Work

Diffusion models. Diffusion probabilistic models (Ho et al., 2020; Song et al., 2021b;a) have emerged as state-of-the-art generative models across image, audio, and video domains. Their success largely stems from stable likelihood training and scalable architectures, with large-scale variants such as Stable Diffusion (Rombach et al., 2022; Podell et al., 2023) demonstrating impressive generative capabilities. Recent advances (Ho & Salimans, 2022; Meng et al., 2021; Zhang et al., 2023; Nichol et al., 2021; Lee et al., 2025; Liu et al., 2024; Sun et al., 2024) have sought to improve controllability by conditioning on textual prompts, sketches, or layout, but aligning with more complex human-defined reward signals remains a challenge.

Reinforcement learning for generative models. RL has been widely adopted in language modeling (Rafailov et al., 2023; Ouyang et al., 2022), where it enables alignment with human preferences through pairwise comparisons. In diffusion models, early efforts adapted policy gradient (Agarwal et al., 2019; Black et al., 2023), but these methods often suffer from high variance and unstable credit assignment across long horizons. Recent works therefore explore preference-based training tailored to diffusion models, such as Diffusion-DPO (Wallace et al., 2023) and D3PO (Yang et al., 2023), which directly optimize model likelihood ratios against a reference. However, these approaches typically equate trajectory likelihoods with terminal-state probabilities, an oversimplification that can misalign the learned distribution with the intended reward structure.

Segment-wise training. To mitigate horizon-related difficulties, several works decompose long diffusion rollouts into shorter segments (Liang et al., 2024; Yang et al., 2023). While this improves stability, segment-level objectives are often designed heuristically, and their consistency with the global reward remains unclear. Our work provides a principled alternative: we ground the segment-wise objective in stochastic optimal control, derive a closed-form expression of the optimal value function, and use it to define preference-consistent objectives for Direct Preference Optimization. This yields both theoretical guarantees and practical improvements.

Stochastic optimal control and continuous-time RL. Our approach is also connected to the literature on SOC and continuous-time RL. Linearly-solvable MDPs (Todorov, 2006; 2009) show that optimal controls can often be expressed via exponentiated value functions and Doob h-transforms, a perspective that closely parallels our formulation. In continuous-time settings, stochastic control theory provides tools such as the Hamilton–Jacobi–Bellman equation and Girsanov's theorem (Fleming & Soner, 2006; Oksendal, 2013), which describe how optimal policies reshape diffusion dynamics. Recent works (Domingo-Enrich et al., 2025) have further connected RL objectives with score-based generative modeling. Our method builds on these insights by importing SOC principles into diffusion preference optimization, establishing both theoretical grounding and practical benefits.

D THE TRANING ALGORITHM OF VALUE-DISTRIBUTION FUNCTION

We have a pratical training algorithm as shown in Algorithm 2.

Algorithm 2 Learning the Value-Distribution via CDF Modeling (CRPS + BCE + TV-Smooth)

Require: Reference sampler $p_{\text{ref}}(\mathbf{x}_1)$; reward $r(\mathbf{x}_1)$; forward noising $\text{Noisify}(\mathbf{x}_1, t)$; CDF model $F_\phi(z \mid \mathbf{x}_t, t)$; grid $\{c_m\}_{m=1}^M$; weights $(\lambda_{\text{crps}}, \lambda_{\text{bce}}, \lambda_{\text{tv}})$; small $\varepsilon > 0$

1: **for** each iteration **do**
2: **Sample terminal images:** $\{\mathbf{x}_1^{(i)}\}_{i=1}^B \sim p_{\text{ref}}(\mathbf{x}_1)$
3: **Score images:** $y^{(i)} \leftarrow r(\mathbf{x}_1^{(i)})$
4: **Sample time and noising:** $t^{(i)} \sim \text{Unif}(\mathcal{T}_T)$, $\mathbf{x}_{t^{(i)}}^{(i)} \leftarrow \text{Noisify}\big(\mathbf{x}_1^{(i)}, t^{(i)}\big)$
5: **Evaluate CDF on grid:** $F_m^{(i)} \leftarrow F_\phi\big(c_m \mid \mathbf{x}_{t^{(i)}}^{(i)}, t^{(i)}\big)$, $m = 1{:}M$
6: **CRPS loss:**

$$\mathcal{L}_{\text{crps}} \;=\; \frac{1}{B}\sum_{i=1}^{B}\sum_{m=1}^{M}\Big(F_m^{(i)} - \mathbf{1}\{c_m \geq y^{(i)}\}\Big)^2$$

7: **CDF binary cross-entropy:**

$$\mathcal{L}_{\text{bce}} \;=\; -\frac{1}{B}\sum_{i=1}^{B}\sum_{m=1}^{M}\Big[\mathbf{1}\{c_m \geq y^{(i)}\}\,\log(F_m^{(i)}+\varepsilon) + \big(1-\mathbf{1}\{c_m \geq y^{(i)}\}\big)\,\log(1-F_m^{(i)}+\varepsilon)\Big]$$

8: **TV smoothness on CDF:**

$$\mathcal{L}_{\text{tv}} \;=\; \frac{1}{B}\sum_{i=1}^{B}\frac{1}{M-1}\sum_{m=1}^{M-1}\Big|F_{m+1}^{(i)} - F_m^{(i)}\Big|$$

9: **Total loss and update:**

$$\mathcal{L} \;=\; \lambda_{\text{crps}}\mathcal{L}_{\text{crps}} + \lambda_{\text{bce}}\mathcal{L}_{\text{bce}} + \lambda_{\text{tv}}\mathcal{L}_{\text{tv}}, \qquad \phi \leftarrow \phi - \eta\nabla_\phi\mathcal{L}$$

10: **end for**
11: **Output:** Trained CDF $F_\phi(z \mid \mathbf{x}_t, t)$.
12: **Post hoc value:**

$$V_\phi^*(\mathbf{x}_t, t, \eta) \;\approx\; \log\sum_{m=1}^{M}\tilde{\pi}_m(\mathbf{x}_t, t)\, e^{\eta c_m},$$

$$\tilde{\pi}_m(\mathbf{x}_t, t) = \frac{\big(F_\phi(c_m \mid \mathbf{x}_t, t) - F_\phi(c_{m-1} \mid \mathbf{x}_t, t)\big)_+ + \varepsilon}{\sum_{\ell=1}^{M}\big(F_\phi(c_\ell \mid \mathbf{x}_t, t) - F_\phi(c_{\ell-1} \mid \mathbf{x}_t, t)\big)_+ + M\varepsilon}.$$

E KEY ANALYSIS RESULTS FROM SOC

E.1 OPTIMAL VALUE FUNCTION

From the SOC formulation, the *value function* is defined as the optimal objective under the controlled reverse-time SDE (3):

$$V(\mathbf{x}, t) \;:=\; \min_{\mathbf{u}}\; \mathbb{E}_{\boldsymbol{x}^{\mathbf{u}}\sim p^{\mathbf{u}}}\left[\int_t^1 \tfrac{1}{2}\|\mathbf{u}(\mathbf{x}_s^{\mathbf{u}}, s)\|^2\,\mathrm{d}s - \eta r(\mathbf{x}_1^{\mathbf{u}})\,\Big|\,\mathbf{x}_t^{\mathbf{u}} = \mathbf{x}\right]. \tag{7}$$

Equivalently, by the linearly-solvable control result (Kappen, 2005), the value function admits a representation with respect to the uncontrolled reference process p_{ref}:

$$V(\mathbf{x}, t) \;=\; -\log\,\mathbb{E}_{\boldsymbol{x}\sim p_{\text{ref}}}\Big[\exp\!\big(\eta r(\mathbf{x}_1)\big)\,\Big|\,\mathbf{x}_t = \mathbf{x}\Big], \tag{8}$$

where the expectation is over trajectories $\boldsymbol{x}_{0:1}$ of the base process.

E.2 OPTIMAL SCORE FUNCTION

The optimal control $\mathbf{u}^*(\mathbf{x}, t)$ can be expressed via the gradient of the value function (Kappen, 2005):

$$\mathbf{u}^*(\mathbf{x}, t) \;=\; -\sigma(t)\, \nabla_{\mathbf{x}} V(\mathbf{x}, t). \tag{9}$$

Then the *optimal score function* is

$$\mathbf{s}^*(\mathbf{x}, t) \;=\; \nabla_{\mathbf{x}} \log p_{\text{ref}}(\mathbf{x}, t) - \nabla_{\mathbf{x}} V(\mathbf{x}, t). \tag{10}$$

It is worth noting that $\mathbf{s}^*(\mathbf{x}, t)$ corresponds to the score of the tilted distribution

$$p^*(\mathbf{x}) = \frac{1}{Z}\, p_{\text{ref}}(\mathbf{x}) \exp(\eta r(\mathbf{x})),$$

where Z is the normalization constant.

F PROOFS

F.1 PROOF OF LEMMA 1

Proof. Fix $(\mathbf{x}_t, t)$ and let $\mathbb{P}^{\text{ref}}_{\mathbf{x}_t,t}$ denote the (path) law on trajectories $(\mathbf{x}_s)_{s\in[t,1]}$ induced by the (reference) reverse-time SDE starting at $\mathbf{x}_t$ at time t. For any admissible control $\mathbf{u}$, let $\mathbb{P}^{\mathbf{u}}_{\mathbf{x}_t,t}$ be the corresponding controlled path law. By Girsanov's theorem (under the usual linear growth/Lipschitz assumptions ensuring absolute continuity), we have

$$D_{\text{KL}}\Big(\mathbb{P}^{\mathbf{u}}_{\mathbf{x}_t,t} \,\big\|\, \mathbb{P}^{\text{ref}}_{\mathbf{x}_t,t}\Big) = \mathbb{E}_{\mathbb{P}^{\mathbf{u}}_{\mathbf{x}_t,t}}\left[\int_t^1 \tfrac{1}{2}\,\|\mathbf{u}(\mathbf{x}_s, s)\|^2\, ds\right].$$

Hence the KL–regularized continuous-time RL objective (Equation (5)) can be written as

$$V^*(\mathbf{x}_t, t, \eta) = \sup_{\mathbf{u}}\Big\{\eta\, \mathbb{E}_{\mathbb{P}^{\mathbf{u}}_{\mathbf{x}_t,t}}\big[r(\mathbf{x}_1)\big] - D_{\text{KL}}\big(\mathbb{P}^{\mathbf{u}}_{\mathbf{x}_t,t} \,\|\, \mathbb{P}^{\text{ref}}_{\mathbf{x}_t,t}\big)\Big\}. \tag{11}$$

According to Equation (8), we have

$$V^*(\mathbf{x}_t, t, \eta) = \mathbb{E}_{\mathbf{x}_1\sim p_{\text{ref}}(\mathbf{x}_1|\mathbf{x}_t,t)}\big[\exp\big(\eta\, r(\mathbf{x}_1)\big)\big],$$

which yields the claimed closed form. The expression depends only on the reward distribution given $\mathbf{x}_t$, so the coefficient $1/\eta$ (KL temperature) can be tuned *post hoc* without retraining. □

F.2 PROOF OF LEMMA 2

Proof. Consider the reverse-time SDE on $[t_k, t_k{+}\Delta]$:

$$d\mathbf{x}_t = \Big(f(\mathbf{x}_t, t) + \sigma(t)^2\, s_\theta(\mathbf{x}_t, t)\Big)\, dt + \sigma(t)\, d\mathbf{B}_t, \qquad \mathbf{x}_{t_k} \text{ given.}$$

Assume the usual Lipschitz and linear-growth conditions so that the Euler–Maruyama (EM) scheme is valid. Over a small step $\Delta = \frac{L}{T}$, the EM update at $(\mathbf{x}_{t_k}, t_k)$ reads

$$\mathbf{x}_{t_k+\Delta} \;\approx\; \mathbf{x}_{t_k} + \Big(f(\mathbf{x}_{t_k}, t_k) + \sigma(t_k)^2\, s_\theta(\mathbf{x}_{t_k}, t_k)\Big)\Delta + \sigma(t_k)\big(\mathbf{B}_{t_k+\Delta} - \mathbf{B}_{t_k}\big).$$

Since $\mathbf{B}_{t_k+\Delta} - \mathbf{B}_{t_k} \sim \mathcal{N}(\mathbf{0},\, \Delta\,\mathbf{I})$ and is independent of $\mathbf{x}_{t_k}$, conditioning on the same start $(\mathbf{x}_{t_k}, t_k)$ implies

$$\mathbf{x}_{t_k+\Delta} \,\big|\, \mathbf{x}_{t_k} \;\approx\; \mathcal{N}\Big(\underbrace{\mathbf{x}_{t_k} + \big(f(\mathbf{x}_{t_k}, t_k) + \sigma(t_k)^2 s_\theta(\mathbf{x}_{t_k}, t_k)\big)\Delta}_{\mu_{t_k}},\; \underbrace{\sigma(t_k)^2\, \Delta\, \mathbf{I}}_{\Sigma_{t_k}}\Big).$$

Standard EM error bounds give $\|\mathbb{E}[\varphi(\mathbf{x}_{t_k+\Delta})] - \mathbb{E}[\varphi(\mathbf{x}^{\text{EM}}_{t_k+\Delta})]\| = \mathcal{O}(\Delta)$ for smooth test functions φ (weak order 1), and $\mathbb{E}\|\mathbf{x}_{t_k+\Delta} - \mathbf{x}^{\text{EM}}_{t_k+\Delta}\|^2 = \mathcal{O}(\Delta)$ (strong order $1/2$), so the Gaussian approximation holds to first order in Δ. □

F.3 Proof of Theorem 3

We decompose the proof into three parts: the characterization of the segment-level optimum, the inductive propagation of optimal marginals, and the base case at $t = 0$.

Define the *desirability* (exponentiated value) at any $(\mathbf{x}, t)$ by

$$z^*(\mathbf{x}, t; \eta) \ := \ \exp\big(V^*(\mathbf{x}, t, \eta)\big) \ = \ \mathbb{E}_{\mathbf{x}_1 \sim p_{\mathrm{ref}}(\cdot|\mathbf{x},t)}\big[\exp\big(\eta\, r(\mathbf{x}_1)\big)\big] \,.$$

By the Markov property of p_{ref} and the law of total expectation, z^* satisfies the *multiplicative Bellman equation* for any intermediate time t and next time $t{+}\Delta$:

$$z^*(\mathbf{x}, t; \eta) \ = \ \mathbb{E}_{\mathbf{x}' \sim p_{\mathrm{ref}}(\cdot|\mathbf{x},t)}\big[z^*(\mathbf{x}', t{+}\Delta; \eta)\big], \qquad V^*(\mathbf{x}, t, \eta) \ = \ \log \mathbb{E}_{\mathbf{x}' \sim p_{\mathrm{ref}}(\cdot|\mathbf{x},t)}\big[e^{V^*(\mathbf{x}', t+\Delta, \eta)}\big]. \tag{12}$$

Lemma 4 (Segment-level VRPO optimum)**.** *Fix a state $(\mathbf{x}, t)$ and a short segment to $t{+}\Delta$. The VRPO optimum with temperature β satisfies*

$$p_\theta(\mathbf{x}' \mid \mathbf{x}, t) \ = \ p_{\mathrm{ref}}(\mathbf{x}' \mid \mathbf{x}, t) \ \exp\Big(\tfrac{1}{\beta} V^*(\mathbf{x}', t{+}\Delta, 1/\beta) - \tfrac{1}{\beta} V^*(\mathbf{x}, t, 1/\beta)\Big). \tag{13}$$

That is, the optimal conditional distribution is the Doob h-transform of the reference dynamics tilted by the terminal reward.

Proof. The VRPO loss with temperature β is

$$\mathcal{L}_{\mathrm{VRPO}} = -\,\mathbb{E}\left[\log \mathrm{sigmoid}\left(\beta \log \frac{p_\theta(\mathbf{x}'^{+} \mid \mathbf{x}, t)}{p_{\mathrm{ref}}(\mathbf{x}'^{+} \mid \mathbf{x}, t)} - \beta \log \frac{p_\theta(\mathbf{x}'^{-} \mid \mathbf{x}, t)}{p_{\mathrm{ref}}(\mathbf{x}'^{-} \mid \mathbf{x}, t)}\right)\right],$$

with preferences $\mathbf{x}'^{+} \succ \mathbf{x}'^{-}$ iff $V^*(\mathbf{x}'^{+}, t{+}\Delta, 1/\beta) > V^*(\mathbf{x}'^{-}, t{+}\Delta, 1/\beta)$. Bradley–Terry analysis implies the optimal ratio has the form

$$\log \frac{p_\theta(\mathbf{x}' \mid \mathbf{x}, t)}{p_{\mathrm{ref}}(\mathbf{x}' \mid \mathbf{x}, t)} = C(\mathbf{x}, t) + \tfrac{1}{\beta} V^*(\mathbf{x}', t{+}\Delta, 1/\beta).$$

Normalizing over $\mathbf{x}'$ and using Equation (12), gives (13). □

Lemma 5 (Inductive propagation of marginals)**.** *Define the tilted marginal*

$$p^*(\mathbf{x}, t) \ \propto \ p_{\mathrm{ref}}(\mathbf{x}, t) \ \exp\big(\tfrac{1}{\beta} V^*(\mathbf{x}, t, 1/\beta)\big). \tag{14}$$

If $p_\theta(\mathbf{x}, t) = p^*(\mathbf{x}, t)$ at time t, then applying the conditional kernel (13) yields $p_\theta(\mathbf{x}', t{+}\Delta) = p^*(\mathbf{x}', t{+}\Delta)$.

Proof. Starting from $p_\theta(\mathbf{x}, t) = p^*(\mathbf{x}, t)$, propagate one segment:

$$\begin{aligned}
p_\theta(\mathbf{x}', t{+}\Delta) &= \int p_\theta(\mathbf{x}' \mid \mathbf{x}, t)\, p_\theta(\mathbf{x}, t)\, d\mathbf{x} \\
&\propto \int p_{\mathrm{ref}}(\mathbf{x}' \mid \mathbf{x}, t) \ \exp\Big(\tfrac{1}{\beta} V^*(\mathbf{x}', t{+}\Delta, 1/\beta) - \tfrac{1}{\beta} V^*(\mathbf{x}, t, 1/\beta)\Big) \\
&\qquad\qquad \cdot p_{\mathrm{ref}}(\mathbf{x}, t) \ \exp\big(\tfrac{1}{\beta} V^*(\mathbf{x}, t, 1/\beta)\big)\, d\mathbf{x} \\
&= \exp\big(\tfrac{1}{\beta} V^*(\mathbf{x}', t{+}\Delta, 1/\beta)\big)\, p_{\mathrm{ref}}(\mathbf{x}', t{+}\Delta) \ \propto \ p^*(\mathbf{x}', t{+}\Delta).
\end{aligned}$$

Thus the tilted form is preserved. □

Lemma 6 (Base case under independence (Domingo-Enrich et al., 2025))**.** *Suppose that under the reference process the initial and terminal states are independent, i.e., $p_{\mathrm{ref}}(\mathbf{x}_1 \mid \mathbf{x}_0, 0) = p_{\mathrm{ref}}(\mathbf{x}_1, 1)$. Then the optimal value satisfies*

$$V^*(\mathbf{x}_0, 0, \eta) = \log \mathbb{E}_{\mathbf{x}_1 \sim p_{\mathrm{ref}}(\cdot, 1)}\big[e^{\eta r(\mathbf{x}_1)}\big] =: C_\eta, \qquad V^*(\mathbf{x}_1, 1, \eta) = \eta\, r(\mathbf{x}_1).$$

In particular, $V^*(\mathbf{x}_0, 0, \eta)$ is a constant (independent of $\mathbf{x}_0$).

Proof. By definition of the value (closed form),

$$V^*(\mathbf{x}_0, 0, \eta) = \log \mathbb{E}_{\mathbf{x}_1 \sim p_{\mathrm{ref}}(\cdot|\mathbf{x}_0,0)}\big[e^{\eta r(\mathbf{x}_1)}\big].$$

Under the independence assumption, $p_{\mathrm{ref}}(\mathbf{x}_1 \mid \mathbf{x}_0, 0) = p_{\mathrm{ref}}(\mathbf{x}_1, 1)$, hence

$$V^*(\mathbf{x}_0, 0, \eta) = \log \mathbb{E}_{\mathbf{x}_1 \sim p_{\mathrm{ref}}(\cdot,1)}\big[e^{\eta r(\mathbf{x}_1)}\big] =: C_\eta,$$

which does not depend on $\mathbf{x}_0$. At the terminal time, the boundary condition gives $V^*(\mathbf{x}_1, 1, \eta) = \eta\, r(\mathbf{x}_1)$. □

Lemma 7 (Base case and terminal marginal)**.** *At $t = 0$, the initial distribution of the learned process coincides with that of the reference process:*

$$p_\theta(\mathbf{x}, 0) \;=\; p_{\mathrm{ref}}(\mathbf{x}, 0). \tag{15}$$

Proof. Since $\mathbf{x}_0$ and $\mathbf{x}_1$ are independent under the reference process, Lemma 6 implies

$$p_\theta(\mathbf{x}, 0) \;\propto\; p_{\mathrm{ref}}(\mathbf{x}, 0)\, \exp(C_\eta),$$

where C_η is a constant independent of $\mathbf{x}$. Normalization then yields (15). □

Theorem 3 (Optimality of VRPO)**.** *Consider a trajectory on $\mathcal{T}_T$ that is decomposed into K short segments, each of length L with $\Delta = \frac{L}{T}$ sufficiently small. If VRPO is applied to segment-level preferences defined by $V_k^*(\mathbf{x}_t, t, \eta)$, then under sufficient model capacity, the resulting model is equivalent to training directly on the optimal target distribution*

$$p(\mathbf{x}_1) \;\propto\; p_{data}(\mathbf{x}_1)\, \exp\!\big(\eta\, r(\mathbf{x}_1)\big).$$

Proof of Theorem 3. Combining Lemmas 4, 5, and 7, we conclude that stage-wise VRPO training yields a terminal distribution

$$p(\mathbf{x}_1) \;\propto\; p_{\mathrm{data}}(\mathbf{x}_1)\, \exp(\eta r(\mathbf{x}_1)),$$

which is exactly the optimal target. Moreover, it is straightforward to verify that the intermediate distributions also coincide. □

Table 4: Results of quantitative comparison with baselines using the preference comparison prompts.

Methods	PickScore	imageReward	Aesthetic score	HPSv2
SDXL	22.2645	0.6738	6.4385	0.2786
SPO	23.3715	1.0234	6.8181	**0.2909**
Ours	**23.6044**	**1.1323**	**6.8622**	0.2903

G Detailed Prompts

Prompts for Figure 1 are provided in Table 5. Prompts for Figure 3 are provided in Table 6.

H MLLM Evaluation Prompts

The prompts used for evaluation with ChatGPT-5 were as follows:

> Please evaluate and compare two images based on the following prompt and three dimensions:
>
> Prompt associated with the images: {prompt}
>
> Evaluation dimensions: 1. Overall performance: Comprehensive assessment of image quality, expressiveness and how well it fulfills the prompt 2. Visual appeal: Evaluation of the aesthetics, color matching and composition of the image 3. Text image alignment: Assessment of how well the image content matches the provided prompt
>
> For each dimension, please clearly judge:

Table 5: Prompts for Figure 1.

Row & Line	Prompt
Row 1, Line 1	A beautiful natural woman.
Row 1, Line 3	A drawing of a funny giraffe eating from a tree.
Row 1, Line 4	Red and white high tops shoes and Bitcoin.
Row 1, Line 5	A golden retriever representing god.
Row 1, Line 6	Throne, dark scene, moonlight.
Row 1, Line 7	Highway to hell that goes into a mountain covered in snow.
Row 2, Line 3	Cute girl, Kyoto animation, 4k, high resolution.
Row 2, Line 4	Corgi with helmet on bicycle.
Row 2, Line 5	Fire sorcerer.
Row 2, Line 6	A huskey playing footbal on the beach.
Row 3, Line 1	Tripod legged robot in a lake.
Row 3, Line 2	A robot that is made out of wood.
Row 3, Line 3	Raccoon with a sombrero riding on a atv with Tequila in hand, photorealism.
Row 3, Line 4	Hello kitty mecha, gears of war, style Artstation, octane render, unreal engine 6, epic game Graphics, Fantasy, cyberpunk, conceptual art, Ray tracing.
Row 3, Line 5	A gloomy rabbit drinks wine.

Table 6: Prompts for Figure 3.

Line	Prompt
Line 1	A beautiful Indian woman by the beach.
Line 2	A man eating a glazed donut and a woman eating a chocolate cake.
Line 3	Beautiful dancing unicorns.
Line 4	A full-body shot of a beautiful woman wearing a dress, with sharp focus on her eyes, depicted in an elegant, intricate style by artgerm, jason chan, and mark hill.
Line 5	Cowby angel standing in a field of flowers.
Line 6	A cake with butterflies and rainbows.
Line 7	A coloring book page of a lotus flower, white clear background.
Line 8	A bee devouring the world.

- Image 1 is better
- Image 2 is better
- Both are equal

Please answer in a clear and concise format, with each dimension on a separate line, for example:

- Overall performance: Image 1 is better
- Visual appeal: Both are equal
- Text image alignment: Image 2 is better

Although we allowed ChatGPT-5 to assign an “equal” score in the overall performance dimension, it never selected this option during evaluation.