

FOCUS:
Optimal Control for Multi-Entity World Modeling in Text-to-Image Generation
Supplementary Material

Supplementary Material Contents

A FOCUS Heuristic	2
B Denoising Diffusion as Flow Matching	3
B.1. Variance-Preserving chain to SDE	3
B.2. Reverse-time dynamics	4
B.3. Time change to Flow Matching	4
B.4. Score relations	4
C Evaluation Setup	4
C.1. Evaluation Sampling Parameters	5
C.2. Evaluation Metrics	5
C.3. Test-Time Control: Deterministic Sampling (ODE)	5
C.4. Test-Time Control: Hyperparameters	6
C.5. Fine-tuning Hyperparameters	6
C.6. Additional Metric: Open-Vocabulary Detection	6
D Human Study	7
D.1. Setup	7
D.2. Elo Rating Computation	7
E StableDiffusion 1.5	8
E.1. Implementation: U-Net Architecture	8
E.2. Empirical Results	8
F Self-Cross Guidance (SCG)	9
F.1. Evaluation in Our Pipeline	9
F.2. Evaluation on the SCG Dataset	9
G Extra Samples	12
G.1. Test-Time Control: Stable Diffusion 3.5	12
G.2. Test-Time Control: FLUX.1 [dev]	13
G.3. Test-Time Control: Stable Diffusion 1.5	14
G.4. Fine-tuned: Stable Diffusion 3.5	15
G.5. Fine-tuned: FLUX.1 [dev]	16
H Detailed Evaluation Results	17
H.1. Evaluation Results for Test-Time Control: Stable Diffusion 3.5	17
H.2. Evaluation Results for Test-Time Control: FLUX.1 [dev]	18
H.3. Evaluation Results for Test-Time Control: Stable Diffusion 1.5	19
H.4. Evaluation Results for Fine-Tuning: Stable Diffusion 3.5	20
H.5. Evaluation Results for Fine-Tuning: FLUX.1 [dev]	23

A. FOCUS Heuristic

In this section we detail FOCUS, our probabilistic attention heuristic used as a running cost for disentanglement. We emphasize three practical design choices: (i) encoding spatial proximity before measuring divergence, (ii) aggregating attention maps prior to scoring, and (iii) omitting an explicit collapse regularizer.

Spatially aware divergence. We promote separation of subjects by maximizing a Jensen–Shannon divergence (JSD) defined over attention distributions. A naïve computation on flattened maps discards locality, allowing distant activations to interact as if adjacent. To preserve spatial structure, we (i) reshape token-embedding maps to the target aspect ratio (height and width of the resulting image), (ii) apply a simple 2D Gaussian smoothing (symmetric kernel of size 3), and only then (iii) flatten for scoring. This encodes proximity into the flattened attention tensor and mitigates grid-like artifacts during optimization.

Block selection and aggregation. Modern T2I backbones follow Diffusion Transformer designs [27]. Rather than computing scores *per block* and averaging their scores which can result in conflicting update directions, we first aggregate attention and then score. Concretely, we average cross-attention maps over all blocks that process text and image tokens *separately*, producing a single map per token and a consistent optimization direction. Blocks that jointly process text and image tokens are excluded from this aggregation for compatibility.

No explicit collapse regularizer. We experimented with an entropy-based regularizer aimed at discouraging overly concentrated (collapsed) attention. Let $H(\mathbf{p}) = -\sum_i \mathbf{p}_i \log \mathbf{p}_i$ denote the Shannon Entropy and $\hat{H}(\mathbf{p}) = H(\mathbf{p}) / \log d \in [0, 1]$ its normalized version, where d is the number of spatial locations. For each subject we form its mixture distribution \mathbf{m}_s and added

$$\gamma_{\text{reg}} \cdot \frac{1}{|S|} \sum_{\mathbf{s} \in S} \left(1 - \hat{H}(\mathbf{m}_s)\right), \quad (20)$$

scaling by $\gamma_{\text{reg}} > 0$ to control its effect. In our experiments, small γ_{reg} made the term largely inactive, while larger γ_{reg} pushed mass away from the subject rather than stabilizing it, see Fig. 5 for an example. We therefore omit this term in FOCUS and rely on the probabilistic objective described above.

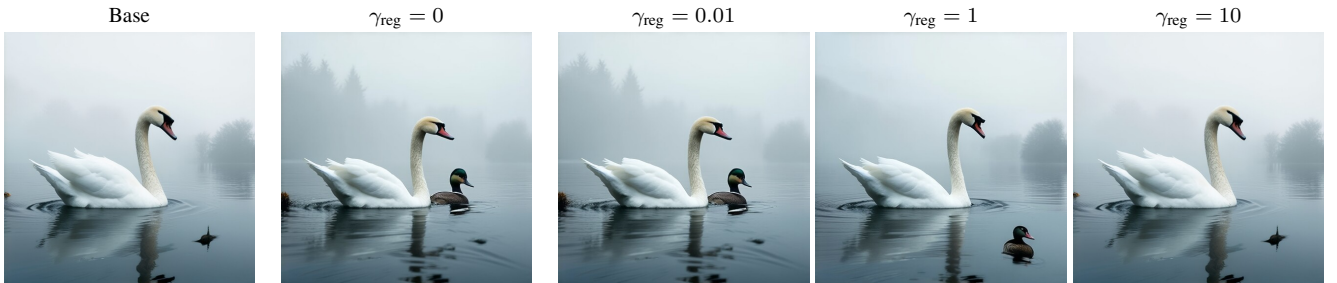


Figure 5. Ablation of regularizer strength γ_{reg} for the test-time controller on Stable Diffusion 3.5.

Lemma A.1 (Upper Bound of Jensen–Shannon Divergence). *Let $P = \{\mathbf{p}^{(1)}, \dots, \mathbf{p}^{(n)}\} \subset \Delta^{d-1}$ be a set of probability distributions. Then, $D_{\text{JS}}(P)$ is upper bounded by $\log n$.*

Proof. Define P as in Theorem A.1, then the JSD is defined as follows:

$$D_{\text{JS}}(P) = \frac{1}{n} \sum_{k=1}^n D_{\text{KL}}\left(\mathbf{p}^{(k)} \parallel \mathbf{m}\right), \quad \mathbf{m} = \frac{1}{n} \sum_{k=1}^n \mathbf{p}^{(k)}.$$

We can upper bound each D_{KL} -term as follows:

$$\begin{aligned}
D_{\text{KL}}(\mathbf{p}^{(k)} \parallel \mathbf{m}) &= \sum_{i=1}^d p_i^{(k)} \log \frac{p_i^{(k)}}{m_i} \\
&= \sum_{i=1}^d p_i^{(k)} \log \frac{p_i^{(k)}}{\frac{1}{n} \sum_{\ell=1}^n p_i^{(\ell)}} \\
&= \sum_{i=1}^d p_i^{(k)} \log \left(n \cdot \frac{p_i^{(k)}}{\sum_{\ell=1}^n p_i^{(\ell)}} \right) \\
&\leq \sum_{i=1}^d p_i^{(k)} \log n \\
&= \log n.
\end{aligned}$$

Plugging this bound back into the definition of the JSD, yields the desired results:

$$\frac{1}{n} \sum_{k=1}^n D_{\text{KL}}(\mathbf{p}^{(k)} \parallel \mathbf{m}) \leq \frac{1}{n} \sum_{k=1}^n \log n = \log n$$

□

Normalization. Because $D_{\text{JS}}(P) \in [0, \log n]$, we use the normalized score $\widehat{D}_{\text{JS}}(P) = D_{\text{JS}}(P)/\log n \in [0, 1]$, which makes values comparable across different set sizes.

B. Denoising Diffusion as Flow Matching

This section makes precise how classical denoising diffusion (score-based) models arise as a special case of the flow-matching (FM) framework. We first derive the continuous-time SDE limit of the variance-preserving (VP) family [12, 32, 33]; after we express reverse-time generation; and finally show an explicit parameterization that uses a diffusion model as an FM velocity field. Analogous statements hold for VE and EDM variants [14, 15, 26].

B.1. Variance-Preserving chain to SDE

Let $X_0 \sim p_{\text{data}}$. The standard K -step VP forward noising chain is

$$X_k = \sqrt{\alpha_k} X_{k-1} + \sqrt{1 - \alpha_k} \epsilon_k, \quad \epsilon_k \sim \mathcal{N}(0, \mathbf{I}), \quad k = 1, \dots, K, \tag{21}$$

where $\alpha_k := 1 - \beta_k \in (0, 1)$ with $\beta_k \in (0, 1)$ typically increasing over k [12]. This yields the closed-form marginal

$$X_k \mid X_0 \sim \mathcal{N}(\sqrt{\bar{\alpha}_k} X_0, (1 - \bar{\alpha}_k) \mathbf{I}), \quad \bar{\alpha}_k := \prod_{i=1}^k \alpha_i. \tag{22}$$

For sufficiently large K , X_K is approximately standard normal [12].

We lift this formulation to continuous time by defining a uniform grid $\tau_k := k/K$ on $[0, 1]$, so every increment is $\Delta\tau = 1/K$. Define a piecewise-constant rate $\beta(\tau)$ via $\beta(\tau) := \beta_k/\Delta\tau$ for $\tau \in [\tau_{k-1}, \tau_k)$. Then by using the first-order Taylor approximation of $\sqrt{1+x}$, we can rewrite $\sqrt{\alpha_k} = \sqrt{1 - \beta_k} \approx 1 - \frac{1}{2}\beta_k + \mathcal{O}(\beta_k^2)$, and obtain

$$\Delta X_k := X_k - X_{k-1} \tag{23}$$

$$= -\frac{1}{2}\beta_k X_{k-1} + \sqrt{\beta_k} \epsilon_k + \mathcal{O}(\beta_k^2) \tag{24}$$

$$= \left(-\frac{1}{2}\beta(\tau_{k-1}) X_{k-1} \right) \Delta\tau + \sqrt{\beta(\tau_{k-1})} \sqrt{\Delta\tau} \epsilon_k + \mathcal{O}\left((\Delta\tau)^{\frac{3}{2}} \right). \tag{25}$$

This is the Euler–Maruyama discretizations of the forward/diffusion VP-SDE:

$$dX_\tau = -\frac{1}{2}\beta(\tau)X_\tau d\tau + \sqrt{\beta(\tau)}dB_\tau, \quad \tau \in [0, 1], \quad (26)$$

and the discrete chain converges to this SDE as $K \rightarrow \infty$. Moreover, the SDE has Gaussian marginals

$$X_\tau | X_0 \sim \mathcal{N}\left(\sqrt{\bar{\alpha}(\tau)}X_0, (1 - \bar{\alpha}(\tau))\mathbf{I}\right), \quad \text{with} \quad \bar{\alpha}(\tau) := \exp\left(-\int_0^\tau \beta(u)du\right), \quad (27)$$

which matches Eq. (22) at the grid points if we choose $\bar{\alpha}(\tau_k) = \bar{\alpha}_k$ [33].

B.2. Reverse-time dynamics

We now reverse time to generate from noise to data. Let $\bar{\tau} = 1 - \tau$ denote the *generative time*. By classical time reversal diffusion [2] the reverse-time process satisfies

$$dX_\tau = \left(-\frac{1}{2}\beta_\tau X_\tau - \beta(\tau)\nabla_X \log p_\tau(X_\tau)\right)d\bar{\tau} + \sqrt{\beta(\tau)}d\bar{B}_\tau, \quad \text{with} \quad d\bar{\tau} = -d\tau, \quad (28)$$

where p_τ are the forward-time marginals and $\nabla_X \log p_\tau$ is the score [33].

In practice, most diffusion architectures parameterize the model via *noise prediction* ϵ_θ [12, 14, 15], which is related to the score by:

$$\nabla_X \log p_\tau(x) = -\frac{\epsilon_\theta(x, \tau)}{\sqrt{1 - \bar{\alpha}(\tau)}}. \quad (29)$$

B.3. Time change to Flow Matching

To embed VP diffusion into FM, we reparameterize time so that FM runs from noise to data, setting $t := 1 - \tau$, which yields the following FM schedules:

$$\alpha_t^{\text{FM}} := \sqrt{\bar{\alpha}(1-t)}, \quad \text{and} \quad \beta_t^{\text{FM}} := \sqrt{1 - \bar{\alpha}(1-t)}. \quad (30)$$

B.4. Score relations

For linear Gaussian reference paths, the score $s(x, t) := \nabla_X \log p_t(x)$ and the FM vector field $v_\theta(x, t)$ are linked by a schedule-dependent affine map [1, 21, 23]:

$$s(x, t) = \frac{1}{\eta_t}\left(v_\theta(x, t) - \kappa_t x\right), \quad \kappa_t := \frac{\dot{\alpha}_t^{\text{FM}}}{\alpha_t^{\text{FM}}}, \quad \eta_t := \beta_t^{\text{FM}}\left(\frac{\dot{\alpha}_t^{\text{FM}}}{\alpha_t^{\text{FM}}}\beta_t^{\text{FM}} - \dot{\beta}_t^{\text{FM}}\right). \quad (31)$$

Combining the noise–score relation with the time change $\tau = 1 - t$ gives:

$$s(x, t) = \nabla_X \log p_t(x) = -\frac{\epsilon_\theta(x, 1-t)}{\beta_t^{\text{FM}}}, \quad (32)$$

since $\beta_t^{\text{FM}} = \sqrt{1 - \bar{\alpha}(1-t)}$. Substituting this into the score–velocity map yields the corresponding FM *velocity prediction* induced by an ϵ -parameterized diffusion model:

$$v_\theta(x, t) = \kappa_t x - \eta_t \frac{\epsilon_\theta(x, 1-t)}{\beta_t^{\text{FM}}}. \quad (33)$$

This identity lets an ϵ -trained diffusion model be used directly as an FM velocity field for the VP-induced schedules above; plugging v_θ into the FM SDE recovers the reverse-time VP sampler (and setting $\sigma \equiv 0$ recovers the probability-flow/DDIM ODE) under the change of variables $t = 1 - \tau$.

C. Evaluation Setup

We detail our evaluation pipeline: sampling parameters, the composite score over metrics, test-time control and fine-tuning hyperparameters, and an auxiliary open-vocabulary detection metric.

C.1. Evaluation Sampling Parameters

For Stable Diffusion 3.5² and FLUX.1 [dev]³, we follow the official sampling recommendations. Unless stated otherwise, we use the deterministic Euler scheduler with 28 inference steps for both models and generate images at 512×512 resolution. The classifier-free guidance scale is set to 4.5 for SD3.5 and 3.5 for FLUX. To ensure consistent extraction of cross-attention maps, we cap the maximum tokenized sequence length at 77 for SD3.5 and 256 for FLUX, and we verify that all prompts in our dataset fall within these limits. Models are loaded and all computations are performed in `bfloat16` to reduce memory usage.

C.2. Evaluation Metrics

To summarize each hyperparameter setting with a single scalar, we macro-average the *relative improvement* over the base model across prompts, seeds, and metrics.

Let $X_{p,s}$ be the image produced by the current setting for prompt $p \in P$ and seed $s \in S$, and let $\hat{X}_{p,s}$ be the corresponding image from the base model. Let \mathcal{M} denote the set of evaluation metrics. Since in our settings all metrics are increasing, the composite score for a hyperparameter setting is the macro-average

$$\frac{1}{|S|} \sum_{s \in S} \frac{1}{|P|} \sum_{p \in P} \frac{1}{|M|} \sum_{m \in M} \frac{m(X_{p,s}) - m(\hat{X}_{p,s})}{m(\hat{X}_{p,s})}, \quad (34)$$

such that a value larger than 0 indicates an average improvement over the base model, while values smaller than 0 indicate degradation.

We conducted 160 test-time control runs and 236 fine-tuning runs, spanning multiple heuristics and hyperparameters. Detailed per-run results are in Section H: test-time control for SD 3.5 in Table 11, for FLUX.1 [dev] in Table 12, and for Stable Diffusion 1.5 in Table 13; fine-tuning results for SD 3.5 in Tables 14 to 16 and for FLUX.1 [dev] in Tables 17 to 19.

C.3. Test-Time Control: Deterministic Sampling (ODE)

Many off-the-shelf T2I models are optimized for deterministic ODE sampling (*i.e.*, $\sigma \equiv 0$). In this setting we view steering as an optimal control problem over the sampler ODE:

$$\min_u \mathbb{E} \left[\int_0^1 \frac{1}{2} \|u(X_t, t)\|_2^2 + f(X_t, t) dt \right] \quad \text{s.t.} \quad \dot{X}_t = v_{\text{base}}(X_t, t) + u(X_t, t), \quad X_0 \sim \pi_0, \quad (35)$$

where v_{base} is the base FM/diffusion velocity parameterization and f is a disentanglement running cost (*e.g.*, FOCUS).

Define the Hamiltonian

$$\mathcal{H}(x, u, a, t) = \frac{1}{2} \|u\|_2^2 + f(x, t) + a^\top (v_{\text{base}}(x, t) + u), \quad (36)$$

with co-state (adjoint) $a(t) \in \mathbb{R}^d$. Since \mathcal{H} is strictly convex in u , the first-order condition gives

$$\nabla_u \mathcal{H} = u + a = 0 \quad \Rightarrow \quad u_t^* = -a(t). \quad (37)$$

Computing u_t^* exactly would require integrating Eq. (9) backward from $t = 1$, which depends on future states and is incompatible with a single forward sampling pass. We therefore use a local approximation analogous to the SDE case: we evaluate gradients at the current state and drop the Jacobian term, $\nabla_x v_{\text{base}}(X_t, t)^\top a(t) \approx 0$. This yields

$$\dot{a}(t) \approx -\nabla_x f(X_t, t), \quad a(1) = 0, \quad (38)$$

so that

$$a(t) \approx \int_t^1 \nabla_x f(X_\tau, \tau) d\tau \approx (1-t) \nabla_x f(X_t, t), \quad (39)$$

²<https://huggingface.co/stabilityai/stable-diffusion-3.5-medium>

³<https://huggingface.co/black-forest-labs/FLUX.1-dev>

where the last step is a left-Riemann approximation (treating the future state locally constant). Substituting Eq. (39) into Eq. (37) gives the instantaneous control

$$u_t^* \approx -(1-t) \nabla_x f(X_t, t). \quad (40)$$

Therefore the controlled ODE uses the guided velocity

$$v_t^* = v_{\text{base}}(X_t, t) + u_t^* \approx v_{\text{base}}(X_t, t) - (1-t) \nabla_x f(X_t, t). \quad (41)$$

With an Euler integrator of step size h , this corresponds to

$$X_{t+h} = X_t + h v_t^* \approx X_t + h v_{\text{base}}(X_t, t) - h(1-t) \nabla_x f(X_t, t), \quad (42)$$

which matches the standard latent-gradient guidance template with an explicit time schedule $\eta(t) = 1 - t$ (up to the global scaling of f by λ). Dropping the Jacobian term $\nabla_x v_{\text{base}}^\top a$ is a common online-control approximation; we empirically find it yields a stable single-pass controller while keeping overhead low [10].

C.4. Test-Time Control: Hyperparameters

In the deterministic (ODE) variant, the single-pass update does not inherit the time-weighting $\frac{1}{2} \sigma_{\text{mem}}^2(t)$ that appears in the SDE case. Since $\sigma_{\text{mem}}(t)$ is large at early times and decays rapidly as $t \rightarrow 1$, we reintroduce this desirable early-strong / late-weak behavior in the ODE setting by reweighting the running cost:

$$f(X_t, t) = \lambda \cdot \sigma_{\text{mem}}^2(t) \cdot \text{Heuristic}(X_t), \quad (43)$$

where $\lambda > 0$ is the earlier introduced hyperparameter to account for different heuristic magnitudes. Throughout our test-time control experiments, we use this time-weighted running cost variant and sweep over $\lambda \in \{0.1, 0.5, 1, 2, 3, 4, 8, 12, 16, 32\}$. Values below 0.1 have negligible effect across heuristics, while values above 32 tend to produce artifacts (over-sharpening, texture noise) or occasional numerical instabilities (NaNs). See Fig. 6 for qualitative trends.



Figure 6. Effect of the control parameter λ on test-time control with Stable Diffusion 3.5.

C.5. Fine-tuning Hyperparameters

We initialize the *memoryless* schedule from each model’s ODE 28-step inference schedule (same time steps), do not use classifier-free guidance, and for **FLUX.1** apply its native guidance scale (not CFG). Following Sec. C.1, we cap tokenized sequence length for cross-attention extraction to 77 (SD 3.5) and 256 (FLUX.1). Models are loaded in `bf16`; forward/backward passes run in BF16 and the final loss reduction is computed in FP32 to avoid numerical issues. To reduce memory, at each iteration we subsample 16 of the 28 steps to be used in our loss calculation. We further use a batch sizes of 5 trajectories for SD 3.5 and 2 trajectories for FLUX.1. We use two small prompt sets: 1 (single prompt: “A horse and a bear”) and 15 (each with two semantically similar subjects). Optimization uses AdamW with a weight decay of 0.01 and $\beta_0 = 0.95$, $\beta_1 = 0.999$. In addition, we also employ `Accelerate` to lower peak memory consumption. Table 5 lists the hyperparameter grids we sweep per heuristic; best settings are **bold**.

C.6. Additional Metric: Open-Vocabulary Detection

As a complementary metric, we assess *subject presence* with OWL-V2 open-vocabulary detection [25]. For each prompt, we pass the subject strings as class queries and count an image as correct if *all* subjects are detected at least once. We report the fraction of images meeting this criterion.

Results for test-time control and fine-tuned models are shown in Tables 6 and 7. Both control algorithms increase subject presence over the base model. However, OWL-V2 is blind to attribute leakage and subject numerosity (it does not verify attributes or counts), so we exclude it from the main evaluation and report it only as a supportive metric here.

Table 5. Hyperparameter grids for fine-tuning; best settings per row in **bold**.

	Heuristic	Lambda λ	Learning rate	Checkpoint	#Prompts
SD 3.5	Attend&Excite	{0.1, 1, 10 }	5e-5	{ 100 , 150 }	1
	CONFORM	{ 0.1 , 1, 10}	5e-5	{ 100 , 150 }	1
	Divide&Bind	{ 0.1 , 1, 10}	5e-5	{ 100 , 150 }	1
	Self-Cross Guidance	{0.1, 1, 10 }	5e-5	{ 100 , 150 }	15
	FOCUS	{0.01, 0.1, 1 , 10, 100}	{1e-4, 5e-5 , 1e-5}	{ 100 , 150, 200}	{ 1 , 15, 150}
FLUX.1	Attend&Excite	{0.1, 1, 10 }	5e-5	{ 200 , 250 }	1
	CONFORM	{0.1, 1, 10 }	5e-5	{ 200 , 250 }	1
	Divide&Bind	{ 0.1 , 1, 10}	5e-5	{ 200 , 250 }	1
	Self-Cross Guidance	{ 0.1 , 1, 10}	5e-5	{ 200 , 250 }	{1, 15 }
	FOCUS	{0.01, 0.1, 1, 10, 100 }	{1e-4, 5e-5 , 1e-5}	{200, 250 , 300}	{1, 15, 150}

Table 6. OWL-V2 subject presence [%] under test-time control. For each heuristic, we report the hyperparameter run with the highest composite score, see Tabs. 11 and 12 for details.

Heuristic	SD3.5	FLUX
Base	69.33%	66.93%
Attend&Excite	72.13%	66.80%
CONFORM	77.20%	67.87%
Divide&Bind	70.80%	68.53%
FOCUS (Ours)	74.27%	68.27%

Table 7. OWL-V2 subject presence [%] under fine-tuned models. For each heuristic, we report the hyperparameter run with the highest composite score, see Tabs. 14 to 19 for details.

Heuristic	SD3.5	FLUX
Base	69.33%	66.93%
Attend&Excite	80.40%	74.93%
CONFORM	77.73%	72.53%
Divide&Bind	73.33%	63.87%
FOCUS (Ours)	78.53%	74.66%

D. Human Study

We test whether automatic metric gains align with human preferences via a prompt-conditioned, pairwise preference study.

D.1. Setup

We evaluate whether metric gains translate to human preferences. Fifty participants each completed 40 prompt-conditioned, pairwise trials, resulting in 2,000 total judgments. In every trial, two images generated from the *same* prompt were shown side by side with the prompt; participants selected the image that better matched the prompt. The instruction shown was:

“Which image renders all subjects of the prompt correctly? If both do an equivalent good job, please pick the one you prefer visually.”

To ensure sufficient rating density, we fixed the sampling seed to 0, yielding one image per method–prompt pair (pool of 150 prompts). Trials were balanced across backbone and setting: SD 3.5 vs. FLUX.1 and test-time control vs. fine-tuning each accounted for one quarter of the comparisons per participant. A screenshot of the interface is shown in Fig. 7.

D.2. Elo Rating Computation

We compute Elo ratings from the pairwise outcomes to obtain an across-method ranking, alongside win rates (fraction of pairwise wins). Elo is initialized at 1500 for all candidates and updated after each comparison with $K=32$. For a candidate A with rating R_A matched against B with R_B , the expected score is

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}, \tag{44}$$

and the update is

$$R'_A = R_A + K (S_A - E_A) \tag{45}$$

where $S_A = 1$ for a win, 0 for a loss, and 0.5 for a draw. Higher Elo indicates stronger preference relative to alternatives. Win rate is reported as the proportion of head-to-head wins.

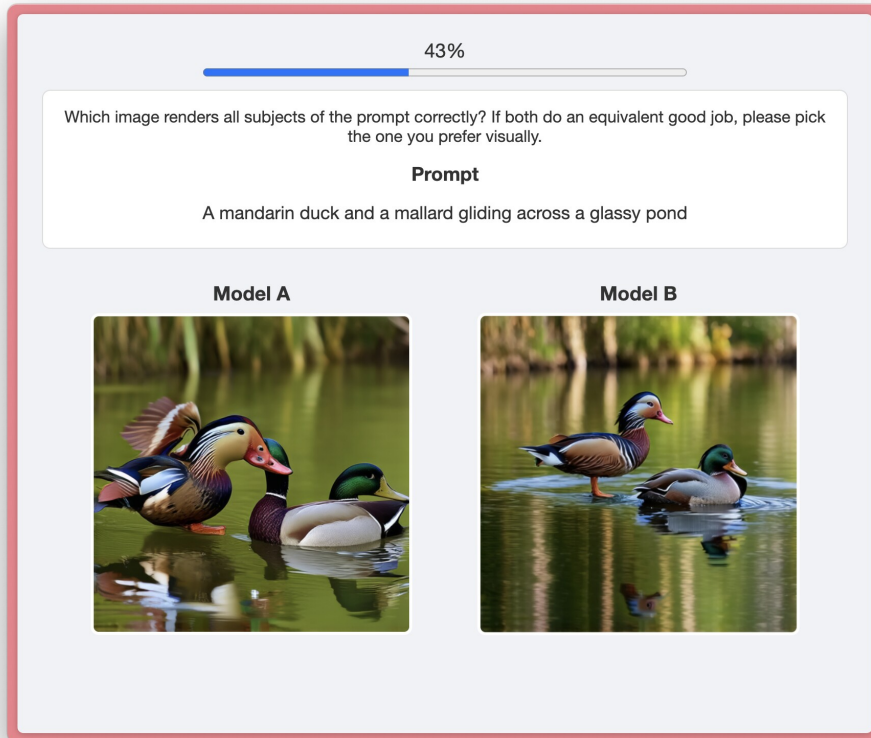


Figure 7. User interface for the prompt-conditioned, pairwise preference study.

E. StableDiffusion 1.5

Although our algorithms are derived for flow matching, Sec. B shows how classical denoising diffusion can be cast in the same framework. To test transferability, we apply the *test-time controller* to *Stable Diffusion 1.5* (SD 1.5)⁴ [31], a U-Net–based denoising diffusion model, and outline the minimal implementation changes below.

E.1. Implementation: U-Net Architecture

Diffusion Transformer (DiT) backbones treat image and text as token sequences processed by stacked Transformer blocks; image–text interaction arises via self-attention of a fixed spatial size. In SD 1.5, based on a U-Net, the latent is down-sampled and upsampled through multiple stages, and text conditioning is injected via cross-attention at several resolutions. Consequently, cross-attention maps have *different spatial sizes* across the network.

To obtain a single subject-specific map per prompt step, we collect *all* cross-attention maps (across down/upsampling and bottleneck), bilinearly resize each to 16×16 , and average them (over heads and layers). This preserves signal from every stage while standardizing spatial shape. Prior works often use only a fixed-resolution subset (*e.g.*, bottleneck) [4, 24]; we found the all-maps aggregation simpler and more comparable across methods, which is the goal of this transfer study.

E.2. Empirical Results

We reuse the main evaluation pipeline for SD 1.5, sweeping 10 values of λ and using the same update schedule as in Equation (43). A qualitative example is shown in Figure 8, with more in Figure 12. Metric results at each heuristic’s best λ appear in Table 8, with the full sweep in Table 13.

Across metrics and prompts, all test-time heuristics improve over the base SD 1.5 model, confirming that our formulation transfers to denoising diffusion. FOCUS achieves the highest composite score and ranks among the top methods on most

⁴<https://huggingface.co/stable-diffusion-v1-5/stable-diffusion-v1-5>

Table 8. Test-time control results for Stable Diffusion 1.5. We report mean \pm std over all prompts and seeds; the top three values per metric are highlighted (gold/silver/bronze). Each method uses the same sampling/evaluation pipeline and its optimal λ .

Heuristic	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
Base	0.3304 \pm 0.03	0.1892 \pm 0.04	0.5433 \pm 0.14	0.5742 \pm 0.09	21.2563 \pm 0.99	-0.0763 \pm 1.00	0.0000 \pm 0.00
Attend&Excite	0.3321 \pm 0.03	0.1964 \pm 0.04	0.5415 \pm 0.15	0.5803 \pm 0.09	21.2082 \pm 1.02	0.0491 \pm 1.05	780.5529 \pm 22303.71
CONFORM	0.3348 \pm 0.03	0.1954 \pm 0.04	0.5449 \pm 0.14	0.5834 \pm 0.09	21.3235 \pm 0.98	0.1137 \pm 0.99	1842.0251 \pm 50959.48
Divide&Bind	0.3350 \pm 0.03	0.1961 \pm 0.04	0.5441 \pm 0.14	0.5817 \pm 0.09	21.3085 \pm 0.99	0.1081 \pm 1.02	1794.9343 \pm 49951.59
FOCUS(Ours)	0.3327 \pm 0.03	0.1952 \pm 0.04	0.5429 \pm 0.14	0.5834 \pm 0.09	21.2977 \pm 0.99	0.1173 \pm 1.02	1862.0311 \pm 51706.75

individual metrics, while preserving the base style and reducing attribute leakage in the qualitative results.



Figure 8. Stable Diffusion 1.5 samples with test-time control. All heuristics shown at their optimal λ . The prompt is “A red fox and an arctic fox sitting side by side in tall grass”.

F. Self-Cross Guidance (SCG)

Self-Cross Guidance (SCG) [29] is, to our knowledge, the only prior heuristic explicitly developed for multi-subject disentanglement on modern DiT-based T2I models (*e.g.*, SD 3/3.5). We therefore include a careful implementation and comparison to FOCUS.

Unlike cross-attention-only objectives, SCG exploits *two* internal signals: (i) text-to-image cross-attention and (ii) image-to-image self-attention. Intuitively, the self-attention term is used to decorrelate image features associated with different subject tokens while the cross-attention term promotes subject binding. This additional signal distinguishes SCG from other heuristics we evaluate.

F.1. Evaluation in Our Pipeline

Under our standard test-time control protocol, SCG marginally exceeds FOCUS on the composite score for SD 3.5. Qualitatively, however, we frequently observe side effects consistent with stronger separation pressure: a tendency toward stylized (cartoon-like) textures and occasional numerosity artifacts (*e.g.*, producing extra instances rather than cleanly separating two subjects); see Figure 9 and additional examples in Figures 10, 11, 13 and 14. This observation is consistent with the limitations mentioned by the authors of SCG. By contrast, FOCUS preserves base style more reliably and maintains subject counts, and achieves higher scores than SCG after fine-tuning.

F.2. Evaluation on the SCG Dataset

For comparability with [29], we also evaluate on their released prompt suite⁵ comprising five subsets: SSD-3 (3 similar subjects; 22 prompts), SSD-2 (2 similar subjects; 31 prompts), Animal-Animal (66 prompts), Animal-Object (144 prompts), and Object-Object (66 prompts). Prompts follow fixed templates such as “a *SUBJECT A* and a *SUBJECT B*” or “a *SUBJECT A* with an *OBJECT B*.”

Test-time control (SD 3.5). Using our standard 10-point λ sweep, both SCG and FOCUS consistently improve over the base model, see Tab. 9. SCG leads on SSD-2, while FOCUS attains higher scores on the remaining subsets, yielding overall comparable performance with a slight average advantage for FOCUS.

⁵<https://github.com/mengtang-lab/selfcross-guidance/blob/main/prompts.txt>

Table 9. Test-time control results on the SCG Prompt Dataset. We report mean \pm std over all prompts and seeds; the top three values per metric are highlighted (gold/silver/bronze). Each method uses the same sampling/evaluation pipeline and its optimal λ .

Dataset	Heuristic	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
SSD-3	Base	0.6494 \pm 0.01	0.5935 \pm 0.01	0.7886 \pm 0.07	0.7996 \pm 0.04	0.2919 \pm 0.11	0.7464 \pm 0.19	0.0000 \pm 0.00
	Self-Cross Guidance	0.6504 \pm 0.01	0.5946 \pm 0.01	0.7969 \pm 0.07	0.8035 \pm 0.04	0.2914 \pm 0.11	0.7689 \pm 0.18	60.6311 \pm 394.49
	FOCUS(Ours)	0.6506 \pm 0.01	0.5957 \pm 0.01	0.7882 \pm 0.08	0.7999 \pm 0.04	0.2928 \pm 0.12	0.7915 \pm 0.15	82.1554 \pm 455.15
SSD-2	Base	0.6477 \pm 0.01	0.5888 \pm 0.02	0.7570 \pm 0.08	0.7904 \pm 0.05	0.2495 \pm 0.10	0.6620 \pm 0.19	0.0000 \pm 0.00
	Self-Cross Guidance	0.6500 \pm 0.01	0.5947 \pm 0.01	0.7598 \pm 0.07	0.7958 \pm 0.05	0.2564 \pm 0.11	0.7277 \pm 0.17	148.2953 \pm 447.50
	FOCUS(Ours)	0.6505 \pm 0.01	0.5958 \pm 0.01	0.7625 \pm 0.08	0.7926 \pm 0.05	0.2531 \pm 0.11	0.7279 \pm 0.17	144.9927 \pm 522.25
Animal—Animal	Base	0.6580 \pm 0.01	0.6036 \pm 0.01	0.8656 \pm 0.07	0.8156 \pm 0.03	0.4239 \pm 0.12	0.9020 \pm 0.13	0.0000 \pm 0.00
	Self-Cross Guidance	0.6583 \pm 0.01	0.6039 \pm 0.01	0.8693 \pm 0.07	0.8170 \pm 0.03	0.4316 \pm 0.11	0.9101 \pm 0.12	35.9864 \pm 343.01
	FOCUS(Ours)	0.6586 \pm 0.01	0.6041 \pm 0.01	0.8667 \pm 0.07	0.8168 \pm 0.03	0.4298 \pm 0.12	0.9177 \pm 0.10	41.9020 \pm 418.06
Animal—Object	Base	0.6710 \pm 0.01	0.6101 \pm 0.01	0.8941 \pm 0.06	0.8415 \pm 0.04	0.5234 \pm 0.15	0.9344 \pm 0.12	0.0000 \pm 0.00
	Self-Cross Guidance	0.6708 \pm 0.01	0.6101 \pm 0.01	0.8954 \pm 0.06	0.8444 \pm 0.03	0.5293 \pm 0.14	0.9361 \pm 0.12	19.2260 \pm 254.76
	FOCUS(Ours)	0.6710 \pm 0.01	0.6101 \pm 0.01	0.8983 \pm 0.05	0.8440 \pm 0.03	0.5308 \pm 0.14	0.9416 \pm 0.11	35.3828 \pm 317.06
Object—Object	Base	0.6754 \pm 0.01	0.6136 \pm 0.02	0.9003 \pm 0.06	0.8519 \pm 0.04	0.5409 \pm 0.16	0.9197 \pm 0.15	0.0000 \pm 0.00
	Self-Cross Guidance	0.6757 \pm 0.01	0.6158 \pm 0.02	0.9053 \pm 0.06	0.8543 \pm 0.04	0.5414 \pm 0.16	0.9320 \pm 0.12	37.8421 \pm 507.36
	FOCUS(Ours)	0.6762 \pm 0.01	0.6154 \pm 0.02	0.9060 \pm 0.06	0.8562 \pm 0.04	0.5529 \pm 0.16	0.9380 \pm 0.11	71.5782 \pm 434.59

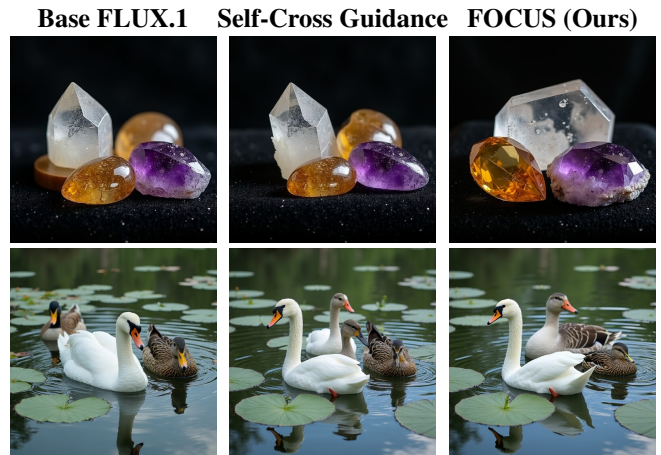
Table 10. Fine-tuned models evaluate on the SCG Prompt Dataset. We report mean \pm std over all prompts and seeds; the top three values per metric are highlighted (gold/silver/bronze). Each method uses the same sampling/evaluation pipeline and its optimal λ .

Dataset	Heuristic	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
SSD-3	Base	0.6494 \pm 0.01	0.5935 \pm 0.01	0.7886 \pm 0.07	0.7996 \pm 0.04	0.2919 \pm 0.11	0.7464 \pm 0.19	0.0000 \pm 0.00
	Self-Cross Guidance	0.6508 \pm 0.01	0.5977 \pm 0.01	0.7931 \pm 0.07	0.8014 \pm 0.04	0.2873 \pm 0.11	0.8139 \pm 0.14	124.4542 \pm 373.52
	FOCUS(Ours)	0.6526 \pm 0.01	0.5987 \pm 0.01	0.7936 \pm 0.07	0.8069 \pm 0.03	0.2957 \pm 0.12	0.8301 \pm 0.12	180.3947 \pm 451.87
SSD-2	Base	0.6477 \pm 0.01	0.5888 \pm 0.02	0.7570 \pm 0.08	0.7904 \pm 0.05	0.2495 \pm 0.10	0.6620 \pm 0.19	0.0000 \pm 0.00
	Self-Cross Guidance	0.6500 \pm 0.01	0.5976 \pm 0.01	0.7666 \pm 0.07	0.8006 \pm 0.05	0.2520 \pm 0.11	0.7578 \pm 0.14	215.1917 \pm 500.96
	FOCUS(Ours)	0.6506 \pm 0.01	0.5988 \pm 0.01	0.7775 \pm 0.07	0.8000 \pm 0.06	0.2523 \pm 0.10	0.7618 \pm 0.13	242.7119 \pm 499.00
Animal—Animal	Base	0.6580 \pm 0.01	0.6036 \pm 0.01	0.8656 \pm 0.07	0.8156 \pm 0.03	0.4239 \pm 0.12	0.9020 \pm 0.13	0.0000 \pm 0.00
	Self-Cross Guidance	0.6583 \pm 0.01	0.6054 \pm 0.01	0.8803 \pm 0.05	0.8236 \pm 0.03	0.4334 \pm 0.11	0.9345 \pm 0.05	111.3952 \pm 372.56
	FOCUS(Ours)	0.6576 \pm 0.01	0.6039 \pm 0.01	0.8754 \pm 0.06	0.8240 \pm 0.03	0.4003 \pm 0.12	0.9230 \pm 0.06	25.9825 \pm 424.15
Animal—Object	Base	0.6710 \pm 0.01	0.6101 \pm 0.01	0.8941 \pm 0.06	0.8415 \pm 0.04	0.5234 \pm 0.15	0.9344 \pm 0.12	0.0000 \pm 0.00
	Self-Cross Guidance	0.6698 \pm 0.01	0.6104 \pm 0.01	0.8972 \pm 0.05	0.8443 \pm 0.03	0.5029 \pm 0.14	0.9509 \pm 0.08	1.3660 \pm 356.49
	FOCUS(Ours)	0.6710 \pm 0.01	0.6110 \pm 0.01	0.8960 \pm 0.06	0.8445 \pm 0.03	0.4888 \pm 0.15	0.9510 \pm 0.08	-20.1645 \pm 382.79
Object—Object	Base	0.6754 \pm 0.01	0.6136 \pm 0.02	0.9003 \pm 0.06	0.8519 \pm 0.04	0.5409 \pm 0.16	0.9197 \pm 0.15	0.0000 \pm 0.00
	Self-Cross Guidance	0.6742 \pm 0.01	0.6149 \pm 0.02	0.9074 \pm 0.06	0.8518 \pm 0.04	0.4943 \pm 0.14	0.9449 \pm 0.10	-23.8872 \pm 463.48
	FOCUS(Ours)	0.6759 \pm 0.01	0.6147 \pm 0.02	0.9111 \pm 0.06	0.8568 \pm 0.03	0.5152 \pm 0.14	0.9494 \pm 0.08	35.6959 \pm 459.01

Fine-tuning generalization. We evaluate the two best SD 3.5 fine-tuned checkpoints trained on our data, and test them *unchanged* on the SCG prompts, see Tab. 10. Both checkpoints generalize and surpass the base model on most subsets. FOCUS underperforms the base model on Animal—Object, whereas SCG is weaker on Object—Object; however, on the general multi-subject splits (SSD-3/SSD-2) both fine-tuned models outperform their test-time counterparts, with SCG leading on Animal—Animal/Animal—Object and FOCUS leading on SSD-3/SSD-2/Object—Object. Additionally, this confirms that our fine-tuned method generalizes well beyond the training dataset, since both checkpoints were achieved with subsets of our own dataset.



(a) Test-time control samples on Stable Diffusion 3.5 for the prompt “A black bear and a brown bear ambling along a riverbank” and “A red car, a blue car, and a green car parked side by side on a city street”.



(b) Test-time control samples on FLUX.1 [dev] for the prompt “A fedora, a beanie, and a baseball cap hanging on a coat rack” and “A swan, a goose, and a duck drifting past lily pads”.



(c) Fine-tuned samples on Stable Diffusion 3.5 for the prompt “A Labrador, a Golden Retriever, and a German Shepherd playing in a backyard” and “A jaguar and a leopard crouching in dense rainforest foliage”.



(d) Fine-tuned samples on FLUX.1 [dev] for the prompt “A macaw, a cockatoo, and an Amazon parrot perched on a jungle vine” and “A red fox and an arctic fox sitting side by side in tall grass”.

Figure 9. Qualitative comparison of Self-Cross Guidance (SCG) and FOCUS. Rows: test-time control (top) and fine-tuned models (bottom). Columns: Stable Diffusion 3.5 (left) and FLUX.1 [dev] (right).

G. Extra Samples

G.1. Test-Time Control: Stable Diffusion 3.5

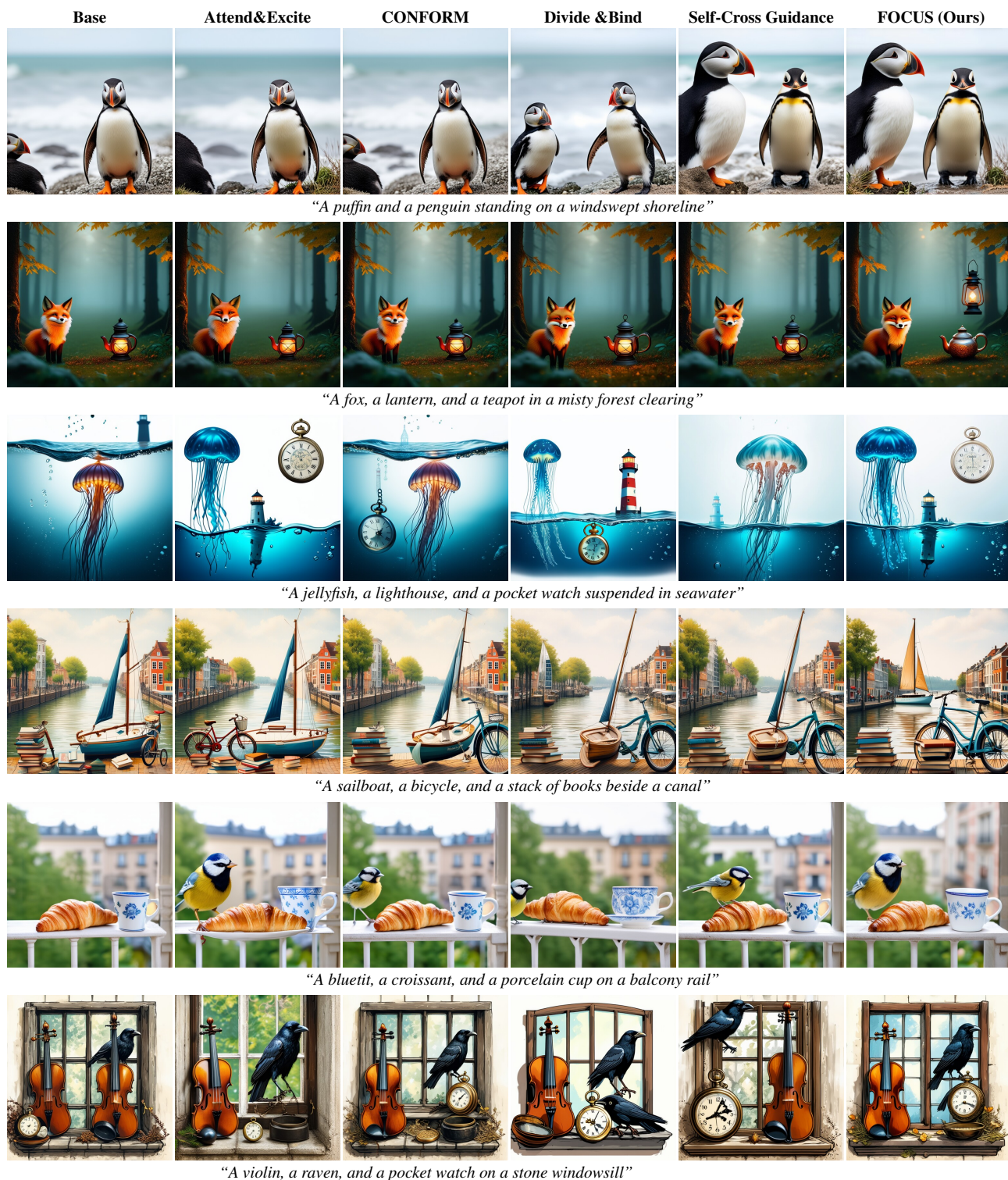


Figure 10. Stable Diffusion 3.5 samples with test-time control. All evaluated heuristics are shown at their optimal λ .

G.2. Test-Time Control: FLUX.1 [dev]

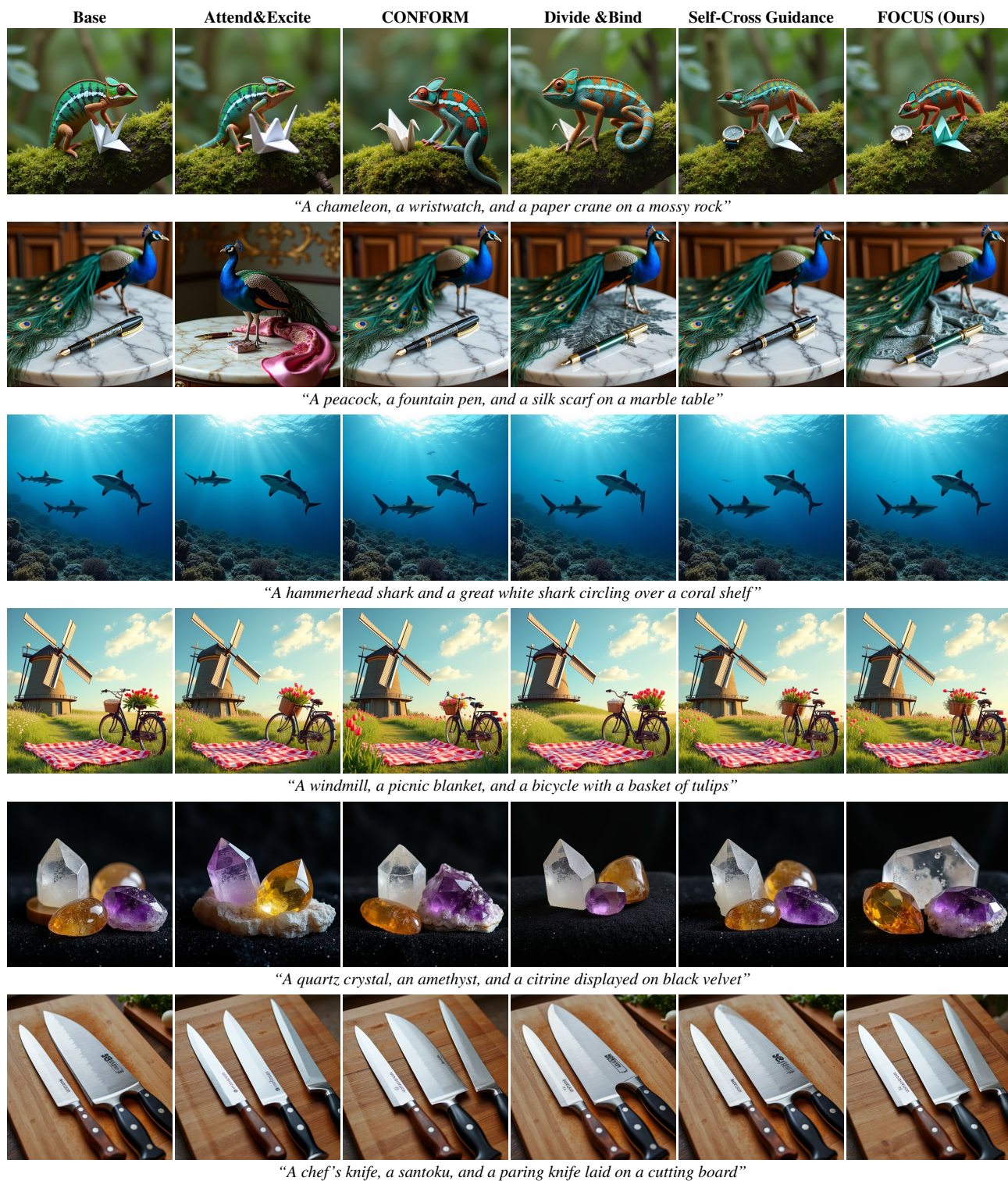


Figure 11. FLUX.1 [dev] samples with test-time control. All evaluated heuristics are shown at their optimal λ .

G.3. Test-Time Control: Stable Diffusion 1.5

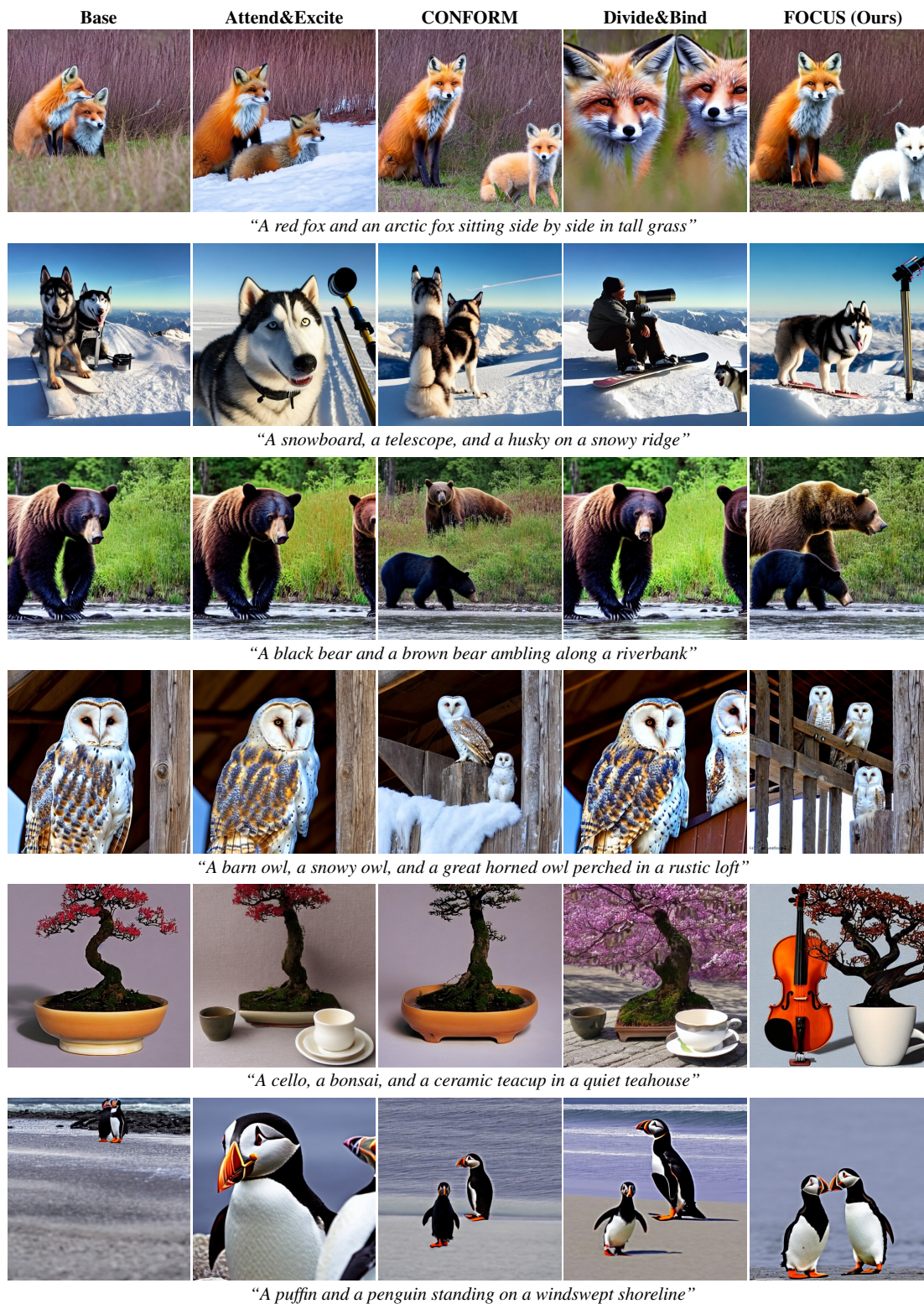


Figure 12. Stable Diffusion 1.5 samples with test-time control. All evaluated heuristics are shown at their optimal λ .

G.4. Fine-tuned: Stable Diffusion 3.5

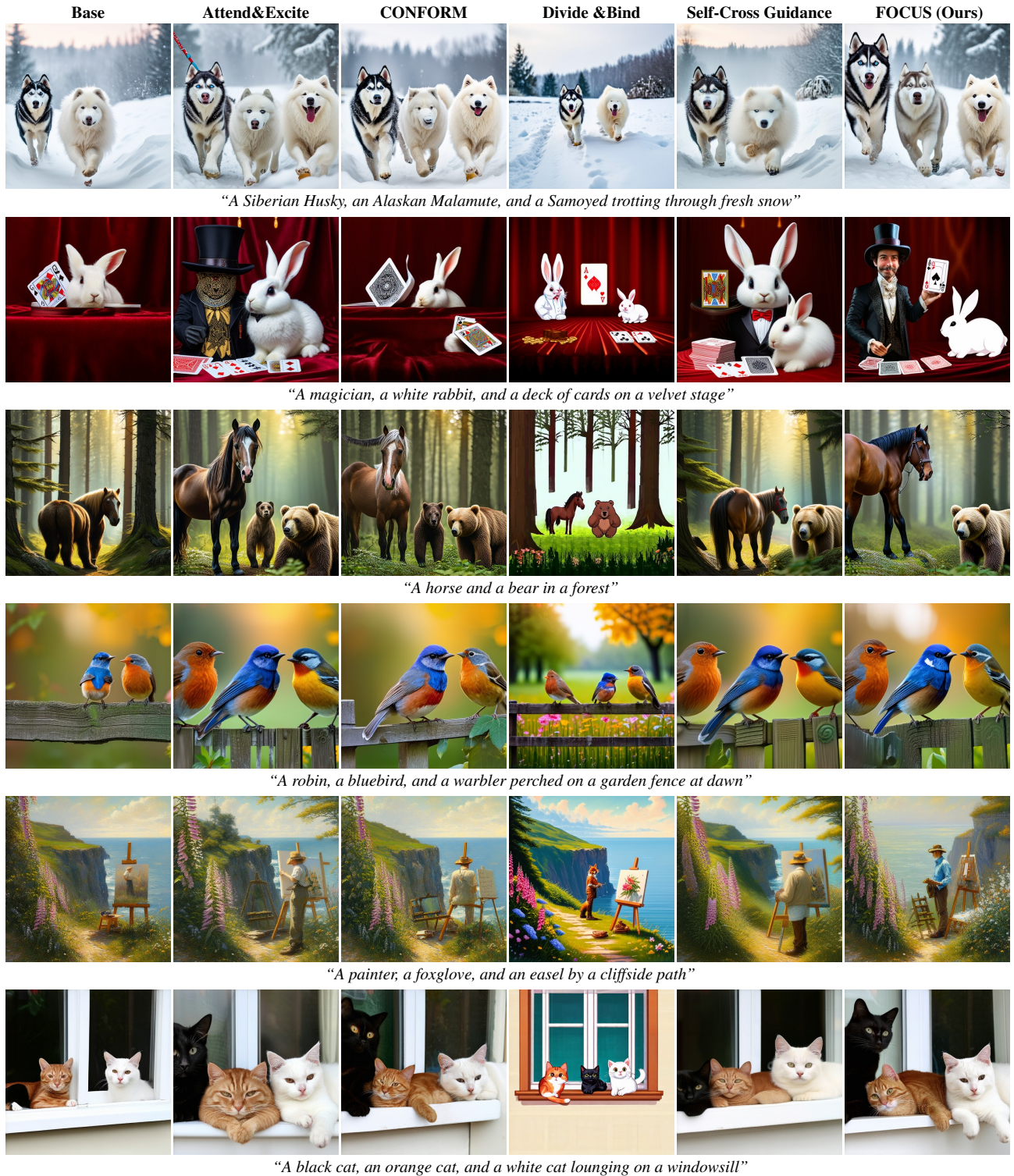


Figure 13. Sample results from Stable Diffusion 3.5 fine-tuned with each heuristic. Prompts were not seen during training to evaluate generalization. All images are generated with identical settings and each heuristic is shown at its optimal trained λ .

G.5. Fine-tuned: FLUX.1 [dev]

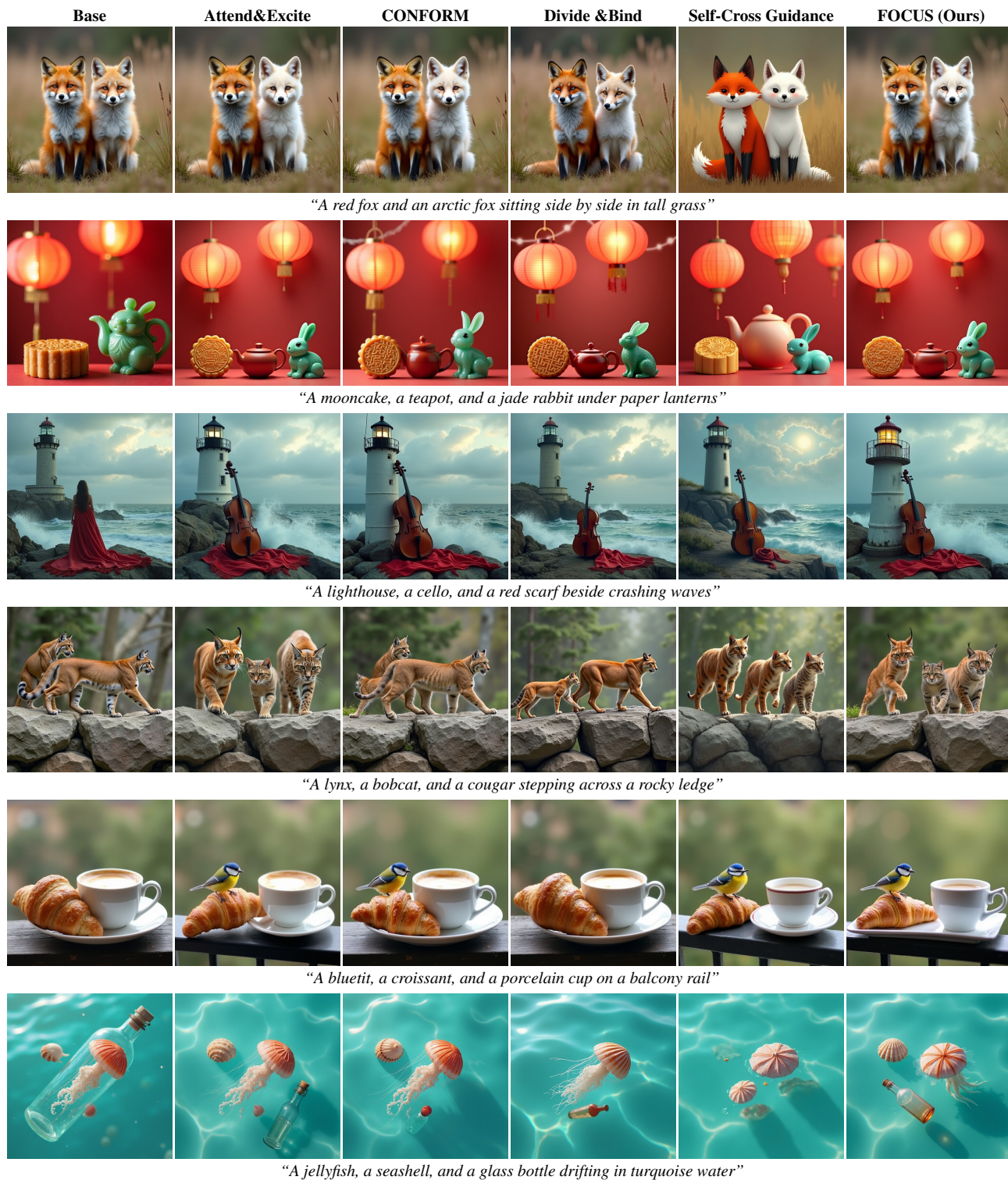


Figure 14. Sample results from FLUX.1 [dev] fine-tuned with each heuristic. Prompts were not seen during training to evaluate generalization. All images are generated with identical settings and each heuristic is shown at its optimal trained λ .

H. Detailed Evaluation Results

H.1. Evaluation Results for Test-Time Control: Stable Diffusion 3.5

Table 11. Test-time control results for each heuristic on Stable Diffusion 3.5. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
	$\lambda = 0$	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5731 \pm 0.15	0.6402 \pm 0.08	22.6940 \pm 0.99	1.3175 \pm 0.68	0.0000 \pm 0.00
Attend&Excite	$\lambda = 0.1$	0.3475 \pm 0.03	0.2308 \pm 0.05	0.5732 \pm 0.15	0.6416 \pm 0.08	22.6990 \pm 0.98	1.3183 \pm 0.68	0.0859 \pm 4.33
	$\lambda = 0.5$	0.3475 \pm 0.03	0.2312 \pm 0.05	0.5733 \pm 0.15	0.6367 \pm 0.08	22.7014 \pm 0.99	1.3304 \pm 0.67	0.0272 \pm 12.73
	$\lambda = 1$	0.3480 \pm 0.03	0.2315 \pm 0.05	0.5738 \pm 0.15	0.6418 \pm 0.08	22.7033 \pm 1.00	1.3323 \pm 0.68	0.8441 \pm 31.03
	$\lambda = 2$	0.3482 \pm 0.03	0.2317 \pm 0.05	0.5739 \pm 0.15	0.6370 \pm 0.08	22.6895 \pm 1.01	1.3364 \pm 0.68	-0.3182 \pm 47.03
	$\lambda = 3$	0.3482 \pm 0.03	0.2323 \pm 0.05	0.5756 \pm 0.15	0.6389 \pm 0.08	22.6951 \pm 1.00	1.3471 \pm 0.67	1.4470 \pm 34.96
	$\lambda = 4$	0.3483 \pm 0.03	0.2319 \pm 0.05	0.5723 \pm 0.15	0.6361 \pm 0.08	22.7126 \pm 1.01	1.3508 \pm 0.67	2.3156 \pm 45.54
	$\lambda = 8$	0.3484 \pm 0.03	0.2326 \pm 0.05	0.5752 \pm 0.15	0.6404 \pm 0.08	22.6950 \pm 1.01	1.3545 \pm 0.66	3.1714 \pm 52.62
	$\lambda = 16$	0.3478 \pm 0.03	0.2322 \pm 0.05	0.5775 \pm 0.14	0.6356 \pm 0.08	22.6601 \pm 1.01	1.3480 \pm 0.68	2.2083 \pm 43.91
	$\lambda = 32$	0.3451 \pm 0.03	0.2294 \pm 0.05	0.5689 \pm 0.15	0.6334 \pm 0.09	22.4442 \pm 1.06	1.2672 \pm 0.73	2.4931 \pm 97.46
CONFORM	$\lambda = 0.1$	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5714 \pm 0.15	0.6376 \pm 0.08	22.6928 \pm 0.99	1.3197 \pm 0.67	0.0078 \pm 13.22
	$\lambda = 0.5$	0.3477 \pm 0.03	0.2317 \pm 0.05	0.5716 \pm 0.15	0.6365 \pm 0.08	22.7023 \pm 0.98	1.3510 \pm 0.65	2.2520 \pm 34.63
	$\lambda = 1$	0.3481 \pm 0.03	0.2323 \pm 0.05	0.5773 \pm 0.15	0.6421 \pm 0.08	22.7188 \pm 0.99	1.3684 \pm 0.64	3.4336 \pm 44.89
	$\lambda = 2$	0.3488 \pm 0.03	0.2335 \pm 0.05	0.5829 \pm 0.14	0.6417 \pm 0.08	22.7255 \pm 1.02	1.3910 \pm 0.63	2.1033 \pm 38.82
	$\lambda = 3$	0.3489 \pm 0.03	0.2348 \pm 0.04	0.5788 \pm 0.14	0.6409 \pm 0.08	22.7446 \pm 1.01	1.4025 \pm 0.62	1.5516 \pm 39.29
	$\lambda = 4$	0.3490 \pm 0.03	0.2351 \pm 0.04	0.5785 \pm 0.15	0.6436 \pm 0.08	22.7400 \pm 1.01	1.4034 \pm 0.62	0.3452 \pm 40.42
	$\lambda = 8$	0.3484 \pm 0.03	0.2348 \pm 0.05	0.5766 \pm 0.15	0.6422 \pm 0.08	22.7414 \pm 1.06	1.4190 \pm 0.63	0.6003 \pm 67.67
	$\lambda = 16$	0.3475 \pm 0.03	0.2339 \pm 0.04	0.5658 \pm 0.15	0.6394 \pm 0.08	22.6293 \pm 1.14	1.3763 \pm 0.69	-2.3354 \pm 65.66
	$\lambda = 32$	0.3414 \pm 0.03	0.2286 \pm 0.05	0.5572 \pm 0.16	0.6257 \pm 0.10	22.2241 \pm 1.33	1.1842 \pm 0.87	-1.7810 \pm 75.92
Divide&Bind	$\lambda = 0.1$	0.3475 \pm 0.03	0.2311 \pm 0.05	0.5723 \pm 0.15	0.6379 \pm 0.08	22.6889 \pm 0.99	1.3203 \pm 0.68	1.2508 \pm 32.26
	$\lambda = 0.5$	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5722 \pm 0.15	0.6386 \pm 0.08	22.6905 \pm 0.98	1.3240 \pm 0.67	0.4329 \pm 11.74
	$\lambda = 1$	0.3476 \pm 0.03	0.2308 \pm 0.05	0.5719 \pm 0.15	0.6391 \pm 0.08	22.6895 \pm 0.98	1.3322 \pm 0.67	0.8767 \pm 15.12
	$\lambda = 2$	0.3474 \pm 0.03	0.2308 \pm 0.05	0.5709 \pm 0.15	0.6360 \pm 0.08	22.6952 \pm 0.98	1.3424 \pm 0.66	1.9613 \pm 33.31
	$\lambda = 3$	0.3485 \pm 0.03	0.2314 \pm 0.05	0.5689 \pm 0.15	0.6410 \pm 0.08	22.7005 \pm 0.96	1.3536 \pm 0.65	1.3799 \pm 26.97
	$\lambda = 4$	0.3486 \pm 0.03	0.2313 \pm 0.05	0.5703 \pm 0.15	0.6415 \pm 0.08	22.7097 \pm 0.98	1.3534 \pm 0.65	2.0489 \pm 31.50
	$\lambda = 8$	0.3482 \pm 0.03	0.2325 \pm 0.05	0.5759 \pm 0.15	0.6371 \pm 0.08	22.7086 \pm 1.02	1.3547 \pm 0.64	3.4455 \pm 41.75
	$\lambda = 16$	0.3489 \pm 0.03	0.2316 \pm 0.05	0.5742 \pm 0.14	0.6399 \pm 0.08	22.6779 \pm 1.03	1.3493 \pm 0.67	3.9373 \pm 77.81
	$\lambda = 32$	0.3470 \pm 0.03	0.2283 \pm 0.05	0.5682 \pm 0.15	0.6358 \pm 0.09	22.4174 \pm 1.08	1.2172 \pm 0.82	-0.6075 \pm 53.02
Self-Cross Guidance	$\lambda = 0.1$	0.3473 \pm 0.03	0.2310 \pm 0.05	0.5695 \pm 0.15	0.6363 \pm 0.08	22.6884 \pm 0.99	1.3199 \pm 0.68	-0.0563 \pm 7.95
	$\lambda = 0.5$	0.3479 \pm 0.03	0.2318 \pm 0.04	0.5705 \pm 0.15	0.6404 \pm 0.08	22.7121 \pm 0.99	1.3442 \pm 0.65	2.3243 \pm 43.65
	$\lambda = 1$	0.3477 \pm 0.03	0.2321 \pm 0.04	0.5695 \pm 0.15	0.6388 \pm 0.08	22.7004 \pm 0.99	1.3527 \pm 0.65	2.8227 \pm 63.03
	$\lambda = 2$	0.3483 \pm 0.03	0.2333 \pm 0.04	0.5747 \pm 0.15	0.6420 \pm 0.08	22.7323 \pm 1.01	1.3640 \pm 0.65	4.8949 \pm 93.18
	$\lambda = 3$	0.3473 \pm 0.03	0.2341 \pm 0.05	0.5711 \pm 0.15	0.6401 \pm 0.08	22.7258 \pm 1.02	1.3744 \pm 0.64	2.9306 \pm 64.24
	$\lambda = 4$	0.3467 \pm 0.03	0.2331 \pm 0.04	0.5757 \pm 0.15	0.6370 \pm 0.08	22.6943 \pm 1.01	1.3610 \pm 0.64	2.1645 \pm 57.01
	$\lambda = 8$	0.3454 \pm 0.03	0.2310 \pm 0.05	0.5688 \pm 0.15	0.6388 \pm 0.09	22.5663 \pm 1.10	1.3042 \pm 0.73	-1.8995 \pm 71.34
	$\lambda = 16$	0.3397 \pm 0.04	0.2231 \pm 0.05	0.5545 \pm 0.15	0.6207 \pm 0.10	22.0941 \pm 1.26	1.0747 \pm 0.97	-10.7541 \pm 139.20
	$\lambda = 32$	0.3199 \pm 0.05	0.1968 \pm 0.06	0.5126 \pm 0.17	0.5592 \pm 0.14	20.9498 \pm 1.56	0.2591 \pm 1.38	-25.9633 \pm 154.64
FOCUS(Ours)	$\lambda = 0.1$	0.3473 \pm 0.03	0.2309 \pm 0.05	0.5705 \pm 0.15	0.6384 \pm 0.08	22.6940 \pm 0.98	1.3186 \pm 0.68	-0.1393 \pm 8.42
	$\lambda = 0.5$	0.3480 \pm 0.03	0.2321 \pm 0.05	0.5742 \pm 0.15	0.6365 \pm 0.08	22.7052 \pm 0.99	1.3448 \pm 0.66	3.2020 \pm 50.28
	$\lambda = 1$	0.3482 \pm 0.03	0.2331 \pm 0.04	0.5761 \pm 0.15	0.6366 \pm 0.08	22.7331 \pm 0.99	1.3703 \pm 0.64	3.6309 \pm 52.76
	$\lambda = 2$	0.3491 \pm 0.03	0.2346 \pm 0.04	0.5778 \pm 0.15	0.6382 \pm 0.08	22.7452 \pm 1.00	1.3868 \pm 0.63	4.1593 \pm 87.43
	$\lambda = 3$	0.3483 \pm 0.03	0.2344 \pm 0.04	0.5751 \pm 0.15	0.6385 \pm 0.08	22.7499 \pm 1.02	1.4003 \pm 0.62	4.2865 \pm 85.86
	$\lambda = 4$	0.3477 \pm 0.03	0.2330 \pm 0.04	0.5749 \pm 0.15	0.6379 \pm 0.08	22.6971 \pm 1.01	1.3842 \pm 0.62	2.5324 \pm 64.04
	$\lambda = 8$	0.3467 \pm 0.03	0.2312 \pm 0.05	0.5692 \pm 0.15	0.6359 \pm 0.09	22.5421 \pm 1.08	1.3280 \pm 0.71	-0.6913 \pm 54.37
	$\lambda = 16$	0.3385 \pm 0.04	0.2233 \pm 0.05	0.5532 \pm 0.15	0.6192 \pm 0.10	21.9730 \pm 1.28	1.0087 \pm 1.02	-9.6793 \pm 107.01
	$\lambda = 32$	0.3191 \pm 0.04	0.1986 \pm 0.05	0.5015 \pm 0.17	0.5607 \pm 0.13	20.9065 \pm 1.52	0.2173 \pm 1.35	-27.0412 \pm 135.97

H.2. Evaluation Results for Test-Time Control: FLUX.1 [dev]

Table 12. Test-time control results for each heuristic on FLUX.1 [Dev]. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
	$\lambda = 0$	0.3449 \pm 0.03	0.2271 \pm 0.05	0.5739 \pm 0.15	0.6300 \pm 0.09	23.4234 \pm 1.03	1.2970 \pm 0.66	0.0000 \pm 0.00
Attend&Excite	$\lambda = 0.1$	0.3450 \pm 0.03	0.2271 \pm 0.05	0.5732 \pm 0.15	0.6294 \pm 0.09	23.4272 \pm 1.03	1.2970 \pm 0.66	0.0806 \pm 8.84
	$\lambda = 0.5$	0.3451 \pm 0.03	0.2272 \pm 0.05	0.5706 \pm 0.15	0.6294 \pm 0.08	23.4223 \pm 1.03	1.2970 \pm 0.66	-0.1954 \pm 14.62
	$\lambda = 1$	0.3451 \pm 0.03	0.2270 \pm 0.05	0.5702 \pm 0.15	0.6328 \pm 0.08	23.4221 \pm 1.04	1.2919 \pm 0.67	0.1986 \pm 14.34
	$\lambda = 2$	0.3450 \pm 0.03	0.2271 \pm 0.05	0.5716 \pm 0.15	0.6317 \pm 0.08	23.4240 \pm 1.05	1.3006 \pm 0.66	0.6700 \pm 19.00
	$\lambda = 3$	0.3449 \pm 0.03	0.2267 \pm 0.05	0.5696 \pm 0.15	0.6294 \pm 0.08	23.4095 \pm 1.03	1.2954 \pm 0.67	-0.4855 \pm 31.68
	$\lambda = 4$	0.3448 \pm 0.03	0.2264 \pm 0.05	0.5703 \pm 0.15	0.6267 \pm 0.09	23.4006 \pm 1.04	1.2706 \pm 0.70	-1.8820 \pm 39.53
	$\lambda = 8$	0.3445 \pm 0.03	0.2261 \pm 0.05	0.5661 \pm 0.15	0.6309 \pm 0.08	23.3708 \pm 1.03	1.2828 \pm 0.67	-1.5221 \pm 46.19
	$\lambda = 16$	0.3430 \pm 0.03	0.2242 \pm 0.05	0.5716 \pm 0.14	0.6304 \pm 0.09	23.2549 \pm 1.11	1.2494 \pm 0.70	1.7595 \pm 67.56
	$\lambda = 32$	0.3363 \pm 0.04	0.2167 \pm 0.05	0.5608 \pm 0.15	0.6131 \pm 0.10	22.6878 \pm 1.47	1.0166 \pm 0.95	-4.0922 \pm 55.70
CONFORM	$\lambda = 0.1$	0.3450 \pm 0.03	0.2271 \pm 0.05	0.5725 \pm 0.15	0.6342 \pm 0.08	23.4247 \pm 1.03	1.3024 \pm 0.66	1.0434 \pm 13.50
	$\lambda = 0.5$	0.3451 \pm 0.03	0.2270 \pm 0.05	0.5724 \pm 0.15	0.6312 \pm 0.09	23.4194 \pm 1.03	1.2978 \pm 0.66	0.5613 \pm 13.56
	$\lambda = 1$	0.3449 \pm 0.03	0.2266 \pm 0.05	0.5694 \pm 0.15	0.6299 \pm 0.08	23.4133 \pm 1.03	1.2874 \pm 0.67	0.2332 \pm 15.24
	$\lambda = 2$	0.3449 \pm 0.03	0.2266 \pm 0.05	0.5679 \pm 0.15	0.6312 \pm 0.09	23.4135 \pm 1.03	1.2864 \pm 0.68	0.3825 \pm 16.61
	$\lambda = 3$	0.3448 \pm 0.03	0.2266 \pm 0.05	0.5689 \pm 0.15	0.6319 \pm 0.08	23.4171 \pm 1.02	1.2842 \pm 0.68	0.2884 \pm 17.10
	$\lambda = 4$	0.3448 \pm 0.03	0.2265 \pm 0.05	0.5699 \pm 0.15	0.6314 \pm 0.08	23.4041 \pm 1.02	1.2824 \pm 0.67	0.6268 \pm 21.74
	$\lambda = 8$	0.3436 \pm 0.03	0.2252 \pm 0.05	0.5726 \pm 0.15	0.6321 \pm 0.09	23.3574 \pm 1.03	1.2461 \pm 0.70	1.5114 \pm 26.28
	$\lambda = 16$	0.3423 \pm 0.03	0.2236 \pm 0.05	0.5730 \pm 0.15	0.6264 \pm 0.09	23.2630 \pm 1.00	1.2100 \pm 0.70	1.1180 \pm 48.29
	$\lambda = 32$	0.3399 \pm 0.03	0.2208 \pm 0.05	0.5660 \pm 0.15	0.6199 \pm 0.09	23.0017 \pm 1.05	1.0908 \pm 0.77	-0.7059 \pm 59.28
Divide&Bind	$\lambda = 0.1$	0.3453 \pm 0.03	0.2272 \pm 0.05	0.5729 \pm 0.15	0.6306 \pm 0.08	23.4246 \pm 1.03	1.2974 \pm 0.67	0.5098 \pm 10.16
	$\lambda = 0.5$	0.3453 \pm 0.03	0.2275 \pm 0.05	0.5699 \pm 0.15	0.6323 \pm 0.08	23.4224 \pm 1.04	1.3003 \pm 0.67	0.5282 \pm 10.97
	$\lambda = 1$	0.3453 \pm 0.03	0.2272 \pm 0.05	0.5740 \pm 0.15	0.6349 \pm 0.08	23.4285 \pm 1.03	1.2964 \pm 0.67	0.8453 \pm 11.31
	$\lambda = 2$	0.3452 \pm 0.03	0.2269 \pm 0.05	0.5717 \pm 0.15	0.6345 \pm 0.08	23.4308 \pm 1.03	1.2892 \pm 0.68	0.5145 \pm 18.48
	$\lambda = 3$	0.3454 \pm 0.03	0.2272 \pm 0.05	0.5742 \pm 0.15	0.6330 \pm 0.09	23.4335 \pm 1.02	1.2876 \pm 0.68	0.5889 \pm 21.60
	$\lambda = 4$	0.3453 \pm 0.03	0.2272 \pm 0.05	0.5722 \pm 0.15	0.6330 \pm 0.08	23.4395 \pm 1.02	1.2939 \pm 0.67	1.6352 \pm 44.00
	$\lambda = 8$	0.3451 \pm 0.03	0.2264 \pm 0.05	0.5709 \pm 0.15	0.6307 \pm 0.08	23.4232 \pm 1.03	1.2865 \pm 0.68	1.1421 \pm 52.58
	$\lambda = 16$	0.3436 \pm 0.03	0.2256 \pm 0.05	0.5725 \pm 0.14	0.6277 \pm 0.09	23.3616 \pm 1.03	1.2593 \pm 0.69	1.1282 \pm 61.75
	$\lambda = 32$	0.3375 \pm 0.04	0.2147 \pm 0.06	0.5511 \pm 0.16	0.6034 \pm 0.12	22.9019 \pm 1.53	1.0226 \pm 0.99	-3.7637 \pm 60.17
Self-Cross Guidance	$\lambda = 0.1$	0.3452 \pm 0.03	0.2272 \pm 0.05	0.5733 \pm 0.15	0.6294 \pm 0.09	23.4286 \pm 1.03	1.2973 \pm 0.67	0.0439 \pm 8.03
	$\lambda = 0.5$	0.3450 \pm 0.03	0.2271 \pm 0.05	0.5795 \pm 0.15	0.6319 \pm 0.08	23.4323 \pm 1.03	1.3026 \pm 0.66	0.4521 \pm 13.95
	$\lambda = 1$	0.3451 \pm 0.03	0.2275 \pm 0.05	0.5734 \pm 0.15	0.6320 \pm 0.09	23.4400 \pm 1.02	1.3072 \pm 0.65	1.9744 \pm 28.28
	$\lambda = 2$	0.3448 \pm 0.03	0.2268 \pm 0.05	0.5722 \pm 0.15	0.6302 \pm 0.09	23.4312 \pm 1.02	1.2963 \pm 0.66	1.9200 \pm 31.86
	$\lambda = 3$	0.3447 \pm 0.03	0.2266 \pm 0.05	0.5722 \pm 0.15	0.6307 \pm 0.09	23.4225 \pm 1.01	1.2903 \pm 0.68	0.9173 \pm 37.31
	$\lambda = 4$	0.3439 \pm 0.03	0.2263 \pm 0.05	0.5733 \pm 0.15	0.6311 \pm 0.08	23.4182 \pm 1.00	1.2872 \pm 0.67	1.0579 \pm 30.83
	$\lambda = 8$	0.3435 \pm 0.03	0.2253 \pm 0.05	0.5670 \pm 0.15	0.6282 \pm 0.09	23.3642 \pm 1.01	1.2435 \pm 0.73	-2.9671 \pm 77.16
	$\lambda = 16$	0.3386 \pm 0.03	0.2192 \pm 0.05	0.5587 \pm 0.16	0.6140 \pm 0.10	23.0488 \pm 1.16	1.1425 \pm 0.80	-4.8268 \pm 67.05
	$\lambda = 32$	0.3042 \pm 0.06	0.1735 \pm 0.08	0.4713 \pm 0.19	0.5246 \pm 0.16	21.0396 \pm 2.13	0.0845 \pm 1.44	-27.1189 \pm 83.54
FOCUS (Ours)	$\lambda = 0.1$	0.3451 \pm 0.03	0.2272 \pm 0.05	0.5725 \pm 0.15	0.6308 \pm 0.08	23.4274 \pm 1.03	1.2982 \pm 0.67	0.2099 \pm 5.98
	$\lambda = 0.5$	0.3450 \pm 0.03	0.2272 \pm 0.05	0.5743 \pm 0.15	0.6352 \pm 0.08	23.4336 \pm 1.02	1.3003 \pm 0.66	0.6785 \pm 19.58
	$\lambda = 1$	0.3450 \pm 0.03	0.2271 \pm 0.05	0.5721 \pm 0.15	0.6332 \pm 0.08	23.4297 \pm 1.03	1.3063 \pm 0.65	0.7895 \pm 26.93
	$\lambda = 2$	0.3446 \pm 0.03	0.2268 \pm 0.05	0.5741 \pm 0.14	0.6326 \pm 0.08	23.4274 \pm 1.02	1.2913 \pm 0.67	1.9712 \pm 31.77
	$\lambda = 3$	0.3443 \pm 0.03	0.2266 \pm 0.05	0.5732 \pm 0.14	0.6303 \pm 0.09	23.4153 \pm 1.01	1.2835 \pm 0.67	1.7279 \pm 33.48
	$\lambda = 4$	0.3443 \pm 0.03	0.2261 \pm 0.05	0.5738 \pm 0.15	0.6270 \pm 0.08	23.3987 \pm 1.00	1.2708 \pm 0.67	0.7435 \pm 34.00
	$\lambda = 8$	0.3439 \pm 0.03	0.2256 \pm 0.05	0.5707 \pm 0.15	0.6314 \pm 0.09	23.3462 \pm 1.01	1.2625 \pm 0.69	0.2671 \pm 51.87
	$\lambda = 16$	0.3395 \pm 0.03	0.2208 \pm 0.05	0.5671 \pm 0.15	0.6211 \pm 0.09	23.0326 \pm 1.13	1.1435 \pm 0.79	-2.7139 \pm 52.43
	$\lambda = 32$	0.3079 \pm 0.05	0.1802 \pm 0.07	0.4914 \pm 0.19	0.5394 \pm 0.16	21.2791 \pm 2.12	0.1595 \pm 1.41	-28.1035 \pm 101.82

H.3. Evaluation Results for Test-Time Control: Stable Diffusion 1.5

Table 13. Test-time control results for each heuristic on Stable Diffusion 1.5. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
	$\lambda = 0$	0.3304 \pm 0.03	0.1892 \pm 0.04	0.5433 \pm 0.14	0.5742 \pm 0.09	21.2563 \pm 0.99	-0.0763 \pm 1.00	0.0000 \pm 0.00
Attend&Excite	$\lambda = 0.1$	0.3309 \pm 0.03	0.1899 \pm 0.04	0.5437 \pm 0.14	0.5736 \pm 0.10	21.2715 \pm 0.99	-0.0718 \pm 1.01	12.2447 \pm 315.89
	$\lambda = 0.5$	0.3310 \pm 0.03	0.1903 \pm 0.04	0.5455 \pm 0.14	0.5744 \pm 0.09	21.2649 \pm 0.99	-0.0621 \pm 1.01	0.5060 \pm 88.13
	$\lambda = 1$	0.3317 \pm 0.03	0.1914 \pm 0.04	0.5390 \pm 0.15	0.5740 \pm 0.10	21.2889 \pm 0.99	-0.0345 \pm 1.02	-10.4363 \pm 225.88
	$\lambda = 2$	0.3322 \pm 0.03	0.1908 \pm 0.04	0.5412 \pm 0.14	0.5767 \pm 0.09	21.2844 \pm 1.03	-0.0429 \pm 1.04	-3.3342 \pm 236.96
	$\lambda = 3$	0.3333 \pm 0.03	0.1928 \pm 0.05	0.5380 \pm 0.15	0.5783 \pm 0.09	21.2833 \pm 1.02	0.0061 \pm 1.05	18.1932 \pm 1061.17
	$\lambda = 4$	0.3328 \pm 0.03	0.1938 \pm 0.04	0.5403 \pm 0.15	0.5771 \pm 0.09	21.2816 \pm 1.02	0.0088 \pm 1.04	149.0053 \pm 4509.00
	$\lambda = 8$	0.3330 \pm 0.03	0.1950 \pm 0.04	0.5470 \pm 0.14	0.5821 \pm 0.09	21.2687 \pm 0.98	0.0523 \pm 1.03	762.6912 \pm 21719.88
	$\lambda = 16$	0.3321 \pm 0.03	0.1964 \pm 0.04	0.5415 \pm 0.15	0.5803 \pm 0.09	21.2082 \pm 1.02	0.0491 \pm 1.05	780.5529 \pm 22303.71
	$\lambda = 32$	0.3273 \pm 0.03	0.1909 \pm 0.04	0.5282 \pm 0.15	0.5685 \pm 0.10	20.9339 \pm 1.08	-0.1262 \pm 1.08	766.6328 \pm 22120.06
CONFORM	$\lambda = 0.1$	0.3313 \pm 0.03	0.1906 \pm 0.04	0.5436 \pm 0.14	0.5750 \pm 0.09	21.2699 \pm 0.99	-0.0641 \pm 1.01	17.5753 \pm 527.26
	$\lambda = 0.5$	0.3329 \pm 0.03	0.1910 \pm 0.04	0.5429 \pm 0.14	0.5795 \pm 0.09	21.3057 \pm 0.99	-0.0044 \pm 1.01	-18.9573 \pm 320.75
	$\lambda = 1$	0.3334 \pm 0.03	0.1932 \pm 0.04	0.5451 \pm 0.14	0.5798 \pm 0.09	21.3268 \pm 0.99	0.0345 \pm 0.99	-10.6165 \pm 94.70
	$\lambda = 2$	0.3348 \pm 0.03	0.1954 \pm 0.04	0.5449 \pm 0.14	0.5834 \pm 0.09	21.3235 \pm 0.98	0.1137 \pm 0.99	1842.0251 \pm 50959.48
	$\lambda = 3$	0.3337 \pm 0.03	0.1949 \pm 0.04	0.5468 \pm 0.15	0.5863 \pm 0.09	21.3070 \pm 0.97	0.1199 \pm 0.98	960.6383 \pm 26741.72
	$\lambda = 4$	0.3321 \pm 0.03	0.1948 \pm 0.04	0.5456 \pm 0.15	0.5848 \pm 0.09	21.2620 \pm 0.96	0.0952 \pm 1.02	1453.7221 \pm 40548.03
	$\lambda = 8$	0.3289 \pm 0.03	0.1914 \pm 0.05	0.5303 \pm 0.15	0.5753 \pm 0.09	21.0058 \pm 1.00	-0.0270 \pm 1.04	-1359.3503 \pm 37540.93
	$\lambda = 16$	0.3163 \pm 0.04	0.1741 \pm 0.05	0.4961 \pm 0.16	0.5374 \pm 0.12	20.3467 \pm 1.23	-0.4981 \pm 1.12	-1811.5944 \pm 49377.24
	$\lambda = 32$	0.2803 \pm 0.05	0.1287 \pm 0.06	0.3859 \pm 0.20	0.4273 \pm 0.16	19.2777 \pm 1.34	-1.4303 \pm 0.97	-2355.0713 \pm 63357.43
Divide&Bind	$\lambda = 0.1$	0.3314 \pm 0.03	0.1902 \pm 0.04	0.5429 \pm 0.14	0.5745 \pm 0.09	21.2699 \pm 0.99	-0.0657 \pm 1.00	-26.2349 \pm 823.24
	$\lambda = 0.5$	0.3314 \pm 0.03	0.1902 \pm 0.04	0.5455 \pm 0.14	0.5767 \pm 0.09	21.2592 \pm 1.00	-0.0238 \pm 1.02	-11.0517 \pm 383.10
	$\lambda = 1$	0.3329 \pm 0.03	0.1911 \pm 0.04	0.5473 \pm 0.14	0.5781 \pm 0.09	21.2959 \pm 0.97	0.0173 \pm 1.02	-5.5725 \pm 159.90
	$\lambda = 2$	0.3331 \pm 0.03	0.1931 \pm 0.04	0.5437 \pm 0.15	0.5792 \pm 0.09	21.2863 \pm 0.99	0.0577 \pm 1.01	-1.0387 \pm 493.01
	$\lambda = 3$	0.3329 \pm 0.03	0.1936 \pm 0.04	0.5371 \pm 0.15	0.5822 \pm 0.09	21.3043 \pm 0.99	0.0578 \pm 1.01	-445.5655 \pm 11554.22
	$\lambda = 4$	0.3350 \pm 0.03	0.1961 \pm 0.04	0.5441 \pm 0.14	0.5817 \pm 0.09	21.3085 \pm 0.99	0.1081 \pm 1.02	1794.9343 \pm 49951.59
	$\lambda = 8$	0.3364 \pm 0.03	0.1985 \pm 0.05	0.5487 \pm 0.14	0.5884 \pm 0.09	21.3312 \pm 1.01	0.1758 \pm 1.04	-2266.3661 \pm 61623.87
	$\lambda = 16$	0.3352 \pm 0.03	0.1970 \pm 0.05	0.5365 \pm 0.15	0.5823 \pm 0.09	21.1888 \pm 1.05	0.1052 \pm 1.08	-2170.6946 \pm 59219.02
	$\lambda = 32$	0.3283 \pm 0.03	0.1917 \pm 0.05	0.5232 \pm 0.14	0.5703 \pm 0.09	20.8602 \pm 1.10	-0.1387 \pm 1.12	-1720.9210 \pm 46815.13
FOCUS(Ours)	$\lambda = 0.1$	0.3309 \pm 0.03	0.1899 \pm 0.04	0.5416 \pm 0.14	0.5758 \pm 0.09	21.2707 \pm 0.99	-0.0601 \pm 1.01	30.7417 \pm 883.40
	$\lambda = 0.5$	0.3316 \pm 0.03	0.1904 \pm 0.04	0.5422 \pm 0.14	0.5785 \pm 0.09	21.2703 \pm 1.00	-0.0456 \pm 1.01	0.6929 \pm 208.23
	$\lambda = 1$	0.3323 \pm 0.03	0.1908 \pm 0.04	0.5449 \pm 0.14	0.5794 \pm 0.09	21.2957 \pm 1.00	-0.0255 \pm 1.00	0.8656 \pm 276.67
	$\lambda = 2$	0.3326 \pm 0.03	0.1924 \pm 0.04	0.5403 \pm 0.15	0.5756 \pm 0.09	21.2865 \pm 0.96	0.0193 \pm 0.99	63.3962 \pm 2171.17
	$\lambda = 3$	0.3334 \pm 0.03	0.1928 \pm 0.04	0.5459 \pm 0.15	0.5791 \pm 0.09	21.3168 \pm 0.97	0.0399 \pm 1.00	37.3881 \pm 1396.57
	$\lambda = 4$	0.3332 \pm 0.03	0.1934 \pm 0.04	0.5422 \pm 0.15	0.5805 \pm 0.09	21.2966 \pm 0.98	0.0813 \pm 1.02	5.4646 \pm 272.59
	$\lambda = 8$	0.3327 \pm 0.03	0.1952 \pm 0.04	0.5429 \pm 0.14	0.5834 \pm 0.09	21.2977 \pm 0.99	0.1173 \pm 1.02	1862.0311 \pm 51706.75
	$\lambda = 16$	0.3320 \pm 0.03	0.1960 \pm 0.05	0.5396 \pm 0.15	0.5803 \pm 0.09	21.1375 \pm 1.01	0.0833 \pm 1.05	1523.1951 \pm 42291.86
	$\lambda = 32$	0.3250 \pm 0.03	0.1899 \pm 0.05	0.5294 \pm 0.15	0.5681 \pm 0.10	20.8645 \pm 1.01	-0.1077 \pm 1.10	698.9492 \pm 19590.17

H.4. Evaluation Results for Fine-Tuning: Stable Diffusion 3.5

Table 14. **Part I:** Fine-tuning results for each heuristic on Stable Diffusion 3.5 across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, $Lambda$ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	N	Lr.	Ckpt.	CLIP I-T\uparrow	SigLIP-2 I-T\uparrow	BLIP T-T\uparrow	Qwen2 T-T\uparrow	PickScore I-T\uparrow	ImgRew I-T\uparrow	Composite\uparrow
	$\lambda = 0$	—	—	—	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5731 \pm 0.15	0.6402 \pm 0.08	22.6940 \pm 0.99	1.3175 \pm 0.68	0.0000 \pm 0.00
Attend&Excite	$\lambda = 0.1$	1	5×10^{-5}	100	0.3469 \pm 0.03	0.2309 \pm 0.04	0.5731 \pm 0.15	0.6385 \pm 0.08	22.7444 \pm 0.99	1.4436 \pm 0.58	4.4775 \pm 100.22
	$\lambda = 0.1$	1	5×10^{-5}	150	0.3448 \pm 0.03	0.2296 \pm 0.05	0.5605 \pm 0.16	0.6343 \pm 0.08	22.6129 \pm 0.99	1.4399 \pm 0.59	3.3100 \pm 88.89
	$\lambda = 1.0$	1	5×10^{-5}	100	0.3463 \pm 0.03	0.2294 \pm 0.05	0.5730 \pm 0.15	0.6398 \pm 0.08	22.8225 \pm 1.00	1.4538 \pm 0.59	5.0286 \pm 106.20
	$\lambda = 1.0$	1	5×10^{-5}	150	0.3427 \pm 0.03	0.2220 \pm 0.05	0.5600 \pm 0.15	0.6325 \pm 0.08	22.5656 \pm 1.02	1.3980 \pm 0.63	0.3568 \pm 65.68
	$\lambda = 10.0$	1	5×10^{-5}	100	0.3469 \pm 0.03	0.2281 \pm 0.04	0.5747 \pm 0.15	0.6425 \pm 0.08	22.8429 \pm 1.01	1.4460 \pm 0.60	5.7181 \pm 121.76
	$\lambda = 10.0$	1	5×10^{-5}	150	0.3417 \pm 0.03	0.2184 \pm 0.05	0.5609 \pm 0.15	0.6291 \pm 0.08	22.4882 \pm 1.03	1.3686 \pm 0.64	-0.7861 \pm 55.86
CONFORM	$\lambda = 0.1$	1	5×10^{-5}	100	0.3478 \pm 0.03	0.2294 \pm 0.05	0.5646 \pm 0.15	0.6393 \pm 0.09	22.5962 \pm 0.99	1.3782 \pm 0.63	3.4583 \pm 105.28
	$\lambda = 0.1$	1	5×10^{-5}	150	0.3475 \pm 0.03	0.2270 \pm 0.05	0.5633 \pm 0.15	0.6348 \pm 0.08	22.5444 \pm 1.02	1.3403 \pm 0.66	2.0584 \pm 102.67
	$\lambda = 1.0$	1	5×10^{-5}	100	0.3476 \pm 0.03	0.2300 \pm 0.04	0.5725 \pm 0.15	0.6431 \pm 0.08	22.6144 \pm 1.01	1.3808 \pm 0.63	3.1881 \pm 98.64
	$\lambda = 1.0$	1	5×10^{-5}	150	0.3473 \pm 0.03	0.2266 \pm 0.05	0.5641 \pm 0.15	0.6393 \pm 0.08	22.5678 \pm 1.01	1.3569 \pm 0.65	1.7289 \pm 98.04
	$\lambda = 10.0$	1	5×10^{-5}	100	0.3479 \pm 0.03	0.2299 \pm 0.05	0.5635 \pm 0.15	0.6407 \pm 0.08	22.6261 \pm 1.00	1.3726 \pm 0.64	2.1542 \pm 103.80
	$\lambda = 10.0$	1	5×10^{-5}	150	0.3474 \pm 0.03	0.2261 \pm 0.05	0.5619 \pm 0.15	0.6401 \pm 0.08	22.5985 \pm 1.00	1.3582 \pm 0.66	2.3135 \pm 108.72
Divide&Bind	$\lambda = 0.1$	1	5×10^{-5}	100	0.3486 \pm 0.03	0.2266 \pm 0.05	0.5870 \pm 0.14	0.6358 \pm 0.08	22.3401 \pm 0.99	1.3524 \pm 0.68	0.8006 \pm 69.71
	$\lambda = 0.1$	1	5×10^{-5}	150	0.3150 \pm 0.04	0.2046 \pm 0.05	0.4834 \pm 0.17	0.5655 \pm 0.12	20.5307 \pm 1.25	0.0025 \pm 1.34	-24.3456 \pm 68.36
	$\lambda = 1.0$	1	5×10^{-5}	100	0.3463 \pm 0.03	0.2237 \pm 0.05	0.5806 \pm 0.15	0.6297 \pm 0.09	22.1119 \pm 1.02	1.2713 \pm 0.74	-0.1148 \pm 70.00
	$\lambda = 1.0$	1	5×10^{-5}	150	0.2666 \pm 0.05	0.1479 \pm 0.06	0.3644 \pm 0.19	0.4511 \pm 0.16	19.0173 \pm 1.48	-1.4507 \pm 1.11	-58.9060 \pm 172.43
	$\lambda = 10.0$	1	5×10^{-5}	100	0.3460 \pm 0.03	0.2230 \pm 0.05	0.5812 \pm 0.15	0.6276 \pm 0.09	22.0908 \pm 1.05	1.2516 \pm 0.77	-2.5875 \pm 56.23
	$\lambda = 10.0$	1	5×10^{-5}	150	0.2849 \pm 0.05	0.1667 \pm 0.06	0.4037 \pm 0.19	0.4939 \pm 0.14	19.4642 \pm 1.46	-1.0902 \pm 1.25	-51.8854 \pm 163.24
Self-Cross Guidance	$\lambda = 0.1$	1	5×10^{-5}	100	0.3437 \pm 0.03	0.2294 \pm 0.05	0.5680 \pm 0.15	0.6302 \pm 0.08	22.4851 \pm 0.96	1.3659 \pm 0.64	1.6630 \pm 119.44
	$\lambda = 0.1$	1	5×10^{-5}	150	0.3387 \pm 0.03	0.2205 \pm 0.05	0.5649 \pm 0.14	0.6208 \pm 0.08	22.0615 \pm 0.92	1.2944 \pm 0.66	1.8345 \pm 117.11
	$\lambda = 0.1$	15	5×10^{-5}	100	0.3450 \pm 0.03	0.2338 \pm 0.04	0.5716 \pm 0.14	0.6337 \pm 0.08	22.6119 \pm 0.97	1.4209 \pm 0.59	3.6476 \pm 84.79
	$\lambda = 0.1$	15	5×10^{-5}	150	0.3376 \pm 0.03	0.2272 \pm 0.04	0.5421 \pm 0.16	0.6146 \pm 0.09	21.8785 \pm 1.02	1.3009 \pm 0.66	-1.2010 \pm 107.81
	$\lambda = 1.0$	1	5×10^{-5}	100	0.3437 \pm 0.03	0.2281 \pm 0.05	0.5805 \pm 0.15	0.6304 \pm 0.08	22.5163 \pm 0.94	1.3521 \pm 0.65	2.6905 \pm 113.90
	$\lambda = 1.0$	1	5×10^{-5}	150	0.3359 \pm 0.03	0.2165 \pm 0.05	0.5632 \pm 0.15	0.6133 \pm 0.09	22.0008 \pm 0.93	1.1914 \pm 0.75	-1.5932 \pm 113.69
	$\lambda = 1.0$	15	5×10^{-5}	100	0.3446 \pm 0.03	0.2339 \pm 0.04	0.5704 \pm 0.15	0.6308 \pm 0.08	22.5440 \pm 0.97	1.4082 \pm 0.60	3.3801 \pm 108.88
	$\lambda = 1.0$	15	5×10^{-5}	150	0.3365 \pm 0.03	0.2257 \pm 0.04	0.5392 \pm 0.16	0.6084 \pm 0.09	21.8108 \pm 1.02	1.2828 \pm 0.66	-1.4363 \pm 87.07
	$\lambda = 10.0$	1	5×10^{-5}	100	0.3432 \pm 0.03	0.2275 \pm 0.04	0.5789 \pm 0.14	0.6284 \pm 0.08	22.4904 \pm 0.95	1.3304 \pm 0.66	2.5557 \pm 124.58
	$\lambda = 10.0$	1	5×10^{-5}	150	0.3324 \pm 0.03	0.2133 \pm 0.05	0.5543 \pm 0.14	0.6059 \pm 0.09	21.8565 \pm 0.92	1.0913 \pm 0.78	-2.3598 \pm 107.26
	$\lambda = 10.0$	15	5×10^{-5}	100	0.3446 \pm 0.03	0.2342 \pm 0.04	0.5715 \pm 0.14	0.6333 \pm 0.08	22.5328 \pm 0.96	1.4097 \pm 0.59	3.7570 \pm 99.01
	$\lambda = 10.0$	15	5×10^{-5}	150	0.3362 \pm 0.03	0.2261 \pm 0.04	0.5518 \pm 0.15	0.6110 \pm 0.09	21.8249 \pm 1.03	1.2696 \pm 0.68	-3.1965 \pm 54.02
FOCUS(Ours)	$\lambda = 0.1$	1	5×10^{-5}	100	0.3493 \pm 0.03	0.2329 \pm 0.04	0.5680 \pm 0.15	0.6425 \pm 0.08	22.6451 \pm 0.98	1.4614 \pm 0.57	4.7817 \pm 108.23
	$\lambda = 0.1$	1	5×10^{-5}	150	0.3488 \pm 0.03	0.2316 \pm 0.04	0.5648 \pm 0.15	0.6386 \pm 0.08	22.4440 \pm 0.97	1.4423 \pm 0.58	4.2309 \pm 94.74
	$\lambda = 0.1$	1	5×10^{-5}	200	0.3484 \pm 0.03	0.2325 \pm 0.04	0.5675 \pm 0.15	0.6371 \pm 0.08	22.4028 \pm 1.00	1.4326 \pm 0.58	4.1562 \pm 96.27
	$\lambda = 0.1$	15	5×10^{-5}	100	0.3473 \pm 0.03	0.2331 \pm 0.04	0.5779 \pm 0.14	0.6369 \pm 0.08	22.6602 \pm 1.01	1.4096 \pm 0.61	3.1910 \pm 99.79
	$\lambda = 0.1$	15	5×10^{-5}	150	0.3336 \pm 0.03	0.2058 \pm 0.05	0.5408 \pm 0.15	0.5832 \pm 0.10	21.4547 \pm 1.08	0.7230 \pm 0.99	-14.7149 \pm 55.03
	$\lambda = 0.1$	15	5×10^{-5}	200	0.3257 \pm 0.04	0.1943 \pm 0.05	0.5235 \pm 0.15	0.5516 \pm 0.12	20.9826 \pm 1.20	0.3691 \pm 1.11	-23.4465 \pm 70.44
	$\lambda = 0.1$	150	5×10^{-5}	100	0.3476 \pm 0.03	0.2290 \pm 0.05	0.5690 \pm 0.14	0.6356 \pm 0.08	22.6842 \pm 1.00	1.3061 \pm 0.67	1.0977 \pm 39.25
	$\lambda = 0.1$	150	5×10^{-5}	150	0.3471 \pm 0.03	0.2285 \pm 0.05	0.5723 \pm 0.14	0.6324 \pm 0.08	22.6144 \pm 1.02	1.2544 \pm 0.69	1.1720 \pm 54.97
	$\lambda = 0.1$	1	1×10^{-4}	100	0.3391 \pm 0.03	0.2204 \pm 0.04	0.5482 \pm 0.15	0.6096 \pm 0.10	21.6393 \pm 1.06	1.1438 \pm 0.75	-3.5377 \pm 79.74
	$\lambda = 0.1$	1	1×10^{-4}	150	0.3412 \pm 0.03	0.2264 \pm 0.04	0.5492 \pm 0.16	0.6201 \pm 0.09	21.8279 \pm 1.04	1.2335 \pm 0.72	1.1393 \pm 107.43
	$\lambda = 0.1$	1	1×10^{-4}	200	0.3422 \pm 0.03	0.2288 \pm 0.04	0.5546 \pm 0.15	0.6209 \pm 0.09	21.8777 \pm 1.02	1.2607 \pm 0.69	1.5729 \pm 99.87
	$\lambda = 0.1$	15	1×10^{-4}	100	0.2620 \pm 0.04	0.1071 \pm 0.05	0.3406 \pm 0.18	0.3395 \pm 0.16	18.8730 \pm 1.41	-1.6188 \pm 0.93	-66.4109 \pm 139.26
	$\lambda = 0.1$	15	1×10^{-4}	150	0.2708 \pm 0.04	0.1161 \pm 0.05	0.3643 \pm 0.18	0.3617 \pm 0.16	19.0972 \pm 1.36	-1.4906 \pm 0.97	-59.5541 \pm 113.78
	$\lambda = 0.1$	15	1×10^{-4}	200	0.2797 \pm 0.04	0.1272 \pm 0.05	0.3878 \pm 0.18	0.3911 \pm 0.16	19.3683 \pm 1.35	-1.3053 \pm 1.03	-54.1921 \pm 113.75
	$\lambda = 0.1$	1	1×10^{-5}	100	0.3477 \pm 0.03	0.2307 \pm 0.05	0.5711 \pm 0.15	0.6390 \pm 0.08	22.6952 \pm 0.99	1.3349 \pm 0.67	0.6719 \pm 22.52
	$\lambda = 0.1$	1	1×10^{-5}	150	0.3477 \pm 0.03	0.2310 \pm 0.05	0.5746 \pm 0.15	0.6375 \pm 0.08	22.6901 \pm 0.98	1.3402 \pm 0.66	1.3884 \pm 38.88
	$\lambda = 0.1$	1	1×10^{-5}	200	0.3476 \pm 0.03	0.2321 \pm 0.05	0.5771 \pm 0.15	0.6387 \pm 0.08	22.7078 \pm 0.97	1.3467 \pm 0.66	1.1869 \pm 36.02
	$\lambda = 0.1$	15	1×10^{-5}	100	0.3475 \pm 0.03	0.2310 \pm 0.05	0.5726 \pm 0.15	0.6385 \pm 0.08	22.6841 \pm 0.98	1.3282 \pm 0.66	0.9204 \pm 16.68
	$\lambda = 0.1$	15	1×10^{-5}	150	0.3477 \pm 0.03	0.2311 \pm 0.05	0.5741 \pm 0.15	0.6370 \pm 0.08	22.6845 \pm 0.98	1.3366 \pm 0.65	1.4782 \pm 31.71

Table 15. **Part II:** Fine-tuning results for each heuristic on Stable Diffusion 3.5 across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, Λ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Λ	N	Lr	$Ckpt.$	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
FOCUS(Ours)	$\lambda = 0.1$	15	1×10^{-5}	200	0.3477 \pm 0.03	0.2314 \pm 0.05	0.5727 \pm 0.15	0.6410 \pm 0.08	22.6869 \pm 0.97	1.3469 \pm 0.65	2.0324 \pm 35.22
	$\lambda = 1.0$	1	5×10^{-5}	100	0.3495 \pm 0.03	0.2331 \pm 0.04	0.5744 \pm 0.15	0.6383 \pm 0.08	22.6445 \pm 0.97	1.4495 \pm 0.58	5.9174 \pm 119.94
	$\lambda = 1.0$	1	5×10^{-5}	150	0.3470 \pm 0.03	0.2273 \pm 0.04	0.5648 \pm 0.15	0.6356 \pm 0.08	22.3249 \pm 0.95	1.3654 \pm 0.64	2.0741 \pm 67.12
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3472 \pm 0.03	0.2244 \pm 0.04	0.5678 \pm 0.14	0.6289 \pm 0.08	22.2283 \pm 0.97	1.3538 \pm 0.64	1.4867 \pm 73.24
	$\lambda = 1.0$	15	5×10^{-5}	100	0.3468 \pm 0.03	0.2324 \pm 0.04	0.5714 \pm 0.14	0.6370 \pm 0.08	22.6213 \pm 0.98	1.4023 \pm 0.63	0.3792 \pm 58.19
	$\lambda = 1.0$	15	5×10^{-5}	150	0.3078 \pm 0.04	0.1722 \pm 0.05	0.4842 \pm 0.17	0.5120 \pm 0.14	20.4852 \pm 1.28	-0.2911 \pm 1.27	-31.6716 \pm 83.22
	$\lambda = 1.0$	15	5×10^{-5}	200	0.2736 \pm 0.04	0.1222 \pm 0.06	0.3826 \pm 0.18	0.3967 \pm 0.16	19.2389 \pm 1.47	-1.3278 \pm 1.10	-58.0962 \pm 127.53
	$\lambda = 1.0$	150	5×10^{-5}	100	0.3478 \pm 0.03	0.2277 \pm 0.05	0.5728 \pm 0.15	0.6329 \pm 0.08	22.6036 \pm 1.04	1.2481 \pm 0.70	0.9205 \pm 42.46
	$\lambda = 1.0$	150	5×10^{-5}	150	0.3472 \pm 0.03	0.2249 \pm 0.05	0.5737 \pm 0.15	0.6344 \pm 0.08	22.5684 \pm 1.04	1.1967 \pm 0.74	0.1973 \pm 43.08
	$\lambda = 1.0$	1	1×10^{-4}	100	0.3065 \pm 0.03	0.1776 \pm 0.05	0.4768 \pm 0.16	0.5068 \pm 0.13	20.2776 \pm 1.26	-0.2760 \pm 1.13	-29.4408 \pm 72.91
	$\lambda = 1.0$	1	1×10^{-4}	150	0.3142 \pm 0.03	0.1867 \pm 0.04	0.4806 \pm 0.17	0.5192 \pm 0.12	20.4988 \pm 1.25	0.0512 \pm 1.15	-24.7687 \pm 66.76
	$\lambda = 1.0$	1	1×10^{-4}	200	0.3137 \pm 0.03	0.1865 \pm 0.05	0.4761 \pm 0.17	0.5126 \pm 0.13	20.4692 \pm 1.26	0.0501 \pm 1.15	-24.2204 \pm 79.37
	$\lambda = 1.0$	15	1×10^{-4}	100	0.2147 \pm 0.03	0.0334 \pm 0.04	0.1803 \pm 0.14	0.2180 \pm 0.10	17.7592 \pm 1.03	-2.2465 \pm 0.17	-92.7828 \pm 186.33
	$\lambda = 1.0$	15	1×10^{-4}	150	0.2069 \pm 0.03	0.0193 \pm 0.03	0.2020 \pm 0.11	0.2002 \pm 0.08	17.2814 \pm 1.06	-2.2737 \pm 0.05	-93.3500 \pm 185.58
	$\lambda = 1.0$	15	1×10^{-4}	200	0.2131 \pm 0.03	0.0299 \pm 0.04	0.2021 \pm 0.11	0.2126 \pm 0.09	17.4684 \pm 1.12	-2.2564 \pm 0.14	-92.0793 \pm 179.00
	$\lambda = 1.0$	1	1×10^{-5}	100	0.3476 \pm 0.03	0.2310 \pm 0.05	0.5742 \pm 0.15	0.6402 \pm 0.08	22.6933 \pm 0.97	1.3300 \pm 0.67	1.7114 \pm 32.33
	$\lambda = 1.0$	1	1×10^{-5}	150	0.3480 \pm 0.03	0.2310 \pm 0.05	0.5717 \pm 0.15	0.6410 \pm 0.08	22.6980 \pm 0.98	1.3446 \pm 0.65	1.9486 \pm 36.14
	$\lambda = 1.0$	1	1×10^{-5}	200	0.3477 \pm 0.03	0.2318 \pm 0.05	0.5777 \pm 0.15	0.6413 \pm 0.08	22.7051 \pm 0.97	1.3551 \pm 0.65	2.5159 \pm 38.56
	$\lambda = 1.0$	15	1×10^{-5}	100	0.3475 \pm 0.03	0.2310 \pm 0.05	0.5741 \pm 0.15	0.6356 \pm 0.08	22.6915 \pm 0.97	1.3302 \pm 0.66	0.5446 \pm 21.15
	$\lambda = 1.0$	15	1×10^{-5}	150	0.3474 \pm 0.03	0.2312 \pm 0.05	0.5712 \pm 0.15	0.6362 \pm 0.08	22.6890 \pm 0.97	1.3385 \pm 0.65	1.5915 \pm 34.08
	$\lambda = 1.0$	15	1×10^{-5}	200	0.3478 \pm 0.03	0.2312 \pm 0.05	0.5729 \pm 0.15	0.6380 \pm 0.08	22.6830 \pm 0.97	1.3486 \pm 0.65	1.2467 \pm 25.99
	$\lambda = 10.0$	1	5×10^{-5}	100	0.3491 \pm 0.03	0.2333 \pm 0.05	0.5735 \pm 0.15	0.6391 \pm 0.08	22.6631 \pm 0.96	1.4431 \pm 0.58	5.4476 \pm 119.33
	$\lambda = 10.0$	1	5×10^{-5}	150	0.3464 \pm 0.03	0.2264 \pm 0.04	0.5675 \pm 0.15	0.6316 \pm 0.08	22.2858 \pm 0.95	1.3305 \pm 0.64	3.3342 \pm 106.31
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3456 \pm 0.03	0.2232 \pm 0.04	0.5676 \pm 0.14	0.6207 \pm 0.09	22.1433 \pm 0.98	1.2824 \pm 0.67	1.5874 \pm 110.54
	$\lambda = 10.0$	15	5×10^{-5}	100	0.3478 \pm 0.03	0.2336 \pm 0.04	0.5717 \pm 0.15	0.6403 \pm 0.08	22.6710 \pm 0.98	1.4244 \pm 0.61	1.8667 \pm 85.84
	$\lambda = 10.0$	15	5×10^{-5}	150	0.3096 \pm 0.04	0.1743 \pm 0.05	0.4942 \pm 0.17	0.5199 \pm 0.13	20.5927 \pm 1.29	-0.3113 \pm 1.26	-30.7088 \pm 101.01
	$\lambda = 10.0$	15	5×10^{-5}	200	0.2627 \pm 0.04	0.1111 \pm 0.06	0.3603 \pm 0.19	0.3780 \pm 0.16	19.0740 \pm 1.48	-1.5467 \pm 1.00	-63.4027 \pm 117.93
	$\lambda = 10.0$	150	5×10^{-5}	100	0.3474 \pm 0.03	0.2272 \pm 0.05	0.5688 \pm 0.15	0.6342 \pm 0.08	22.5742 \pm 1.03	1.2309 \pm 0.73	-1.6672 \pm 27.83
	$\lambda = 10.0$	150	5×10^{-5}	150	0.3455 \pm 0.03	0.2228 \pm 0.05	0.5698 \pm 0.15	0.6285 \pm 0.08	22.4433 \pm 1.04	1.1040 \pm 0.78	-2.3377 \pm 46.06
	$\lambda = 10.0$	1	1×10^{-4}	100	0.2913 \pm 0.04	0.1564 \pm 0.05	0.4332 \pm 0.19	0.4746 \pm 0.15	19.9207 \pm 1.49	-0.8172 \pm 1.23	-44.9893 \pm 83.31
	$\lambda = 10.0$	1	1×10^{-4}	150	0.2772 \pm 0.04	0.1342 \pm 0.06	0.3906 \pm 0.19	0.4192 \pm 0.16	19.3314 \pm 1.57	-1.1540 \pm 1.18	-54.3090 \pm 113.29
	$\lambda = 10.0$	1	1×10^{-4}	200	0.2662 \pm 0.04	0.1207 \pm 0.06	0.3473 \pm 0.19	0.3856 \pm 0.17	18.9562 \pm 1.50	-1.4391 \pm 1.01	-58.3965 \pm 102.75
	$\lambda = 10.0$	15	1×10^{-4}	100	0.2040 \pm 0.03	0.0153 \pm 0.03	0.1327 \pm 0.11	0.1954 \pm 0.08	17.6465 \pm 0.97	-2.2711 \pm 0.05	-97.2981 \pm 187.52
	$\lambda = 10.0$	15	1×10^{-4}	150	0.1938 \pm 0.03	-0.0073 \pm 0.04	0.1764 \pm 0.10	0.1902 \pm 0.07	17.2757 \pm 1.00	-2.2764 \pm 0.03	-97.9584 \pm 188.44
	$\lambda = 10.0$	15	1×10^{-4}	200	0.1971 \pm 0.03	0.0010 \pm 0.04	0.1759 \pm 0.10	0.1891 \pm 0.07	17.2862 \pm 1.03	-2.2753 \pm 0.04	-96.9842 \pm 189.32
	$\lambda = 10.0$	1	1×10^{-5}	100	0.3475 \pm 0.03	0.2310 \pm 0.05	0.5762 \pm 0.15	0.6405 \pm 0.08	22.6962 \pm 0.99	1.3370 \pm 0.67	0.7585 \pm 20.75
	$\lambda = 10.0$	1	1×10^{-5}	150	0.3477 \pm 0.03	0.2313 \pm 0.05	0.5718 \pm 0.15	0.6406 \pm 0.08	22.6990 \pm 0.98	1.3438 \pm 0.66	1.0141 \pm 33.63
	$\lambda = 10.0$	1	1×10^{-5}	200	0.3479 \pm 0.03	0.2320 \pm 0.05	0.5732 \pm 0.15	0.6390 \pm 0.09	22.7105 \pm 0.98	1.3536 \pm 0.66	0.4088 \pm 44.30
	$\lambda = 10.0$	15	1×10^{-5}	100	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5725 \pm 0.15	0.6387 \pm 0.08	22.6884 \pm 0.98	1.3258 \pm 0.66	1.5721 \pm 25.88
	$\lambda = 10.0$	15	1×10^{-5}	150	0.3478 \pm 0.03	0.2312 \pm 0.05	0.5697 \pm 0.15	0.6346 \pm 0.08	22.6800 \pm 0.98	1.3369 \pm 0.65	1.8735 \pm 28.13
$\lambda = 10.0$	15	1×10^{-5}	200	0.3477 \pm 0.03	0.2314 \pm 0.05	0.5717 \pm 0.15	0.6374 \pm 0.08	22.6846 \pm 0.98	1.3458 \pm 0.65	2.2173 \pm 35.69	
$\lambda = 0.01$	1	5×10^{-5}	100	0.3498 \pm 0.03	0.2346 \pm 0.05	0.5731 \pm 0.15	0.6387 \pm 0.08	22.6310 \pm 0.97	1.4409 \pm 0.60	4.7726 \pm 115.00	
$\lambda = 0.01$	1	5×10^{-5}	150	0.3501 \pm 0.03	0.2353 \pm 0.04	0.5663 \pm 0.15	0.6390 \pm 0.08	22.4459 \pm 0.99	1.4257 \pm 0.59	4.6688 \pm 103.96	
$\lambda = 0.01$	1	5×10^{-5}	200	0.3505 \pm 0.03	0.2350 \pm 0.04	0.5678 \pm 0.15	0.6381 \pm 0.08	22.4483 \pm 0.96	1.4258 \pm 0.59	5.4173 \pm 115.75	
$\lambda = 0.01$	15	5×10^{-5}	100	0.3483 \pm 0.03	0.2328 \pm 0.04	0.5776 \pm 0.15	0.6392 \pm 0.08	22.6599 \pm 0.99	1.4011 \pm 0.63	3.1612 \pm 77.83	
$\lambda = 0.01$	15	5×10^{-5}	150	0.3426 \pm 0.03	0.2260 \pm 0.04	0.5609 \pm 0.15	0.6262 \pm 0.09	22.0914 \pm 1.00	1.2267 \pm 0.74	-2.0737 \pm 47.33	
$\lambda = 0.01$	15	5×10^{-5}	200	0.3395 \pm 0.03	0.2221 \pm 0.04	0.5477 \pm 0.15	0.6060 \pm 0.10	21.7539 \pm 1.07	1.0078 \pm 0.88	-5.0309 \pm 70.66	
$\lambda = 0.01$	1	1×10^{-4}	100	0.3498 \pm 0.03	0.2352 \pm 0.04	0.5670 \pm 0.15	0.6386 \pm 0.08	22.3060 \pm 0.97	1.4112 \pm 0.59	4.6317 \pm 113.89	
$\lambda = 0.01$	1	1×10^{-4}	150	0.3505 \pm 0.03	0.2355 \pm 0.04	0.5621 \pm 0.16	0.6387 \pm 0.08	22.3979 \pm 1.00	1.4106 \pm 0.59	3.9209 \pm 111.97	
$\lambda = 0.01$	1	1×10^{-4}	200	0.3513 \pm 0.03	0.2367 \pm 0.04	0.5675 \pm 0.15	0.6394 \pm 0.08	22.4452 \pm 1.01	1.4244 \pm 0.58	5.8620 \pm 111.67	

Table 16. **Part III:** Fine-tuning results for each heuristic on Stable Diffusion 3.5 across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, Λ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Λ	N	Lr	$Ckpt.$	CLIP I-T \uparrow	SigLIP-2 I-T \uparrow	BLIP T-T \uparrow	Qwen2 T-T \uparrow	PickScore I-T \uparrow	ImgRew I-T \uparrow	Composite \uparrow
	$\lambda = 0.01$	15	1×10^{-4}	100	0.3295 \pm 0.03	0.2067 \pm 0.04	0.5226 \pm 0.16	0.5732 \pm 0.11	21.0679 \pm 1.10	0.4918 \pm 1.04	-14.2396 \pm 55.79
	$\lambda = 0.01$	15	1×10^{-4}	150	0.3291 \pm 0.03	0.2060 \pm 0.04	0.5401 \pm 0.15	0.5741 \pm 0.10	21.1208 \pm 1.10	0.4378 \pm 1.05	-13.8534 \pm 56.44
	$\lambda = 0.01$	15	1×10^{-4}	200	0.3347 \pm 0.03	0.2130 \pm 0.04	0.5484 \pm 0.15	0.5851 \pm 0.10	21.3853 \pm 1.06	0.6568 \pm 0.99	-9.7173 \pm 56.37
	$\lambda = 0.01$	1	1×10^{-5}	100	0.3477 \pm 0.03	0.2307 \pm 0.05	0.5773 \pm 0.15	0.6398 \pm 0.08	22.6876 \pm 0.98	1.3314 \pm 0.67	1.9296 \pm 34.01
	$\lambda = 0.01$	1	1×10^{-5}	150	0.3477 \pm 0.03	0.2312 \pm 0.05	0.5745 \pm 0.15	0.6407 \pm 0.08	22.6982 \pm 0.98	1.3428 \pm 0.65	1.7057 \pm 35.15
	$\lambda = 0.01$	1	1×10^{-5}	200	0.3476 \pm 0.03	0.2316 \pm 0.05	0.5730 \pm 0.15	0.6397 \pm 0.08	22.6979 \pm 0.98	1.3515 \pm 0.65	2.5361 \pm 44.72
	$\lambda = 0.01$	15	1×10^{-5}	100	0.3474 \pm 0.03	0.2309 \pm 0.05	0.5703 \pm 0.15	0.6370 \pm 0.08	22.6969 \pm 0.97	1.3279 \pm 0.66	0.3887 \pm 18.03
	$\lambda = 0.01$	15	1×10^{-5}	150	0.3476 \pm 0.03	0.2307 \pm 0.05	0.5681 \pm 0.15	0.6374 \pm 0.08	22.6854 \pm 0.98	1.3354 \pm 0.66	1.6879 \pm 28.53
	$\lambda = 0.01$	15	1×10^{-5}	200	0.3479 \pm 0.03	0.2311 \pm 0.05	0.5708 \pm 0.15	0.6360 \pm 0.08	22.6812 \pm 0.98	1.3432 \pm 0.65	2.0125 \pm 33.64
	$\lambda = 100.0$	1	5×10^{-5}	100	0.3490 \pm 0.03	0.2325 \pm 0.04	0.5743 \pm 0.15	0.6396 \pm 0.08	22.6442 \pm 0.95	1.4427 \pm 0.58	4.9582 \pm 93.26
	$\lambda = 100.0$	1	5×10^{-5}	150	0.3466 \pm 0.03	0.2266 \pm 0.04	0.5723 \pm 0.14	0.6347 \pm 0.08	22.3274 \pm 0.96	1.3414 \pm 0.63	2.1010 \pm 81.88
	$\lambda = 100.0$	1	5×10^{-5}	200	0.3461 \pm 0.03	0.2238 \pm 0.04	0.5724 \pm 0.14	0.6241 \pm 0.09	22.2135 \pm 0.99	1.3255 \pm 0.63	0.0713 \pm 74.41
	$\lambda = 100.0$	15	5×10^{-5}	100	0.3471 \pm 0.03	0.2327 \pm 0.04	0.5706 \pm 0.15	0.6427 \pm 0.08	22.6551 \pm 0.99	1.4199 \pm 0.61	1.8248 \pm 75.46
	$\lambda = 100.0$	15	5×10^{-5}	150	0.3083 \pm 0.04	0.1729 \pm 0.06	0.4849 \pm 0.17	0.5152 \pm 0.13	20.5624 \pm 1.32	-0.3139 \pm 1.29	-30.8746 \pm 99.11
	$\lambda = 100.0$	15	5×10^{-5}	200	0.2585 \pm 0.04	0.1043 \pm 0.06	0.3359 \pm 0.19	0.3592 \pm 0.16	18.9506 \pm 1.49	-1.6260 \pm 0.98	-68.4099 \pm 155.77
	$\lambda = 100.0$	150	5×10^{-5}	100	0.3474 \pm 0.03	0.2275 \pm 0.05	0.5730 \pm 0.15	0.6364 \pm 0.08	22.6208 \pm 1.01	1.2632 \pm 0.71	-0.3567 \pm 42.43
	$\lambda = 100.0$	150	5×10^{-5}	150	0.3470 \pm 0.03	0.2252 \pm 0.05	0.5727 \pm 0.15	0.6324 \pm 0.08	22.5804 \pm 1.04	1.2064 \pm 0.74	1.6821 \pm 53.85
	$\lambda = 100.0$	1	1×10^{-4}	100	0.2814 \pm 0.04	0.1448 \pm 0.06	0.4088 \pm 0.20	0.4481 \pm 0.15	19.6822 \pm 1.44	-1.0771 \pm 1.16	-50.1942 \pm 100.52
	$\lambda = 100.0$	1	1×10^{-4}	150	0.2659 \pm 0.04	0.1165 \pm 0.06	0.3517 \pm 0.19	0.3850 \pm 0.16	19.0181 \pm 1.45	-1.4932 \pm 1.00	-63.5734 \pm 140.11
	$\lambda = 100.0$	1	1×10^{-4}	200	0.2508 \pm 0.04	0.0927 \pm 0.06	0.3014 \pm 0.18	0.3296 \pm 0.16	18.4957 \pm 1.35	-1.8228 \pm 0.74	-74.6304 \pm 178.87
	$\lambda = 100.0$	15	1×10^{-4}	100	0.2128 \pm 0.03	0.0260 \pm 0.04	0.1500 \pm 0.13	0.2069 \pm 0.08	17.8226 \pm 0.98	-2.2638 \pm 0.06	-95.1106 \pm 187.77
	$\lambda = 100.0$	15	1×10^{-4}	150	0.2015 \pm 0.03	0.0159 \pm 0.03	0.1562 \pm 0.11	0.1921 \pm 0.07	17.6473 \pm 0.96	-2.2729 \pm 0.06	-96.4041 \pm 188.52
	$\lambda = 100.0$	15	1×10^{-4}	200	0.1958 \pm 0.03	0.0064 \pm 0.04	0.1798 \pm 0.10	0.1898 \pm 0.07	17.5559 \pm 0.97	-2.2738 \pm 0.05	-96.7049 \pm 188.32
	$\lambda = 100.0$	1	1×10^{-5}	100	0.3477 \pm 0.03	0.2309 \pm 0.05	0.5741 \pm 0.15	0.6385 \pm 0.08	22.6962 \pm 0.99	1.3383 \pm 0.66	1.2236 \pm 23.02
	$\lambda = 100.0$	1	1×10^{-5}	150	0.3479 \pm 0.03	0.2310 \pm 0.05	0.5717 \pm 0.15	0.6398 \pm 0.08	22.6947 \pm 0.98	1.3430 \pm 0.66	1.2978 \pm 38.02
	$\lambda = 100.0$	1	1×10^{-5}	200	0.3478 \pm 0.03	0.2318 \pm 0.05	0.5760 \pm 0.15	0.6409 \pm 0.08	22.6959 \pm 0.98	1.3473 \pm 0.65	-0.2384 \pm 39.03
	$\lambda = 100.0$	15	1×10^{-5}	100	0.3476 \pm 0.03	0.2310 \pm 0.05	0.5704 \pm 0.15	0.6386 \pm 0.08	22.6874 \pm 0.98	1.3318 \pm 0.66	1.7875 \pm 27.87
	$\lambda = 100.0$	15	1×10^{-5}	150	0.3476 \pm 0.03	0.2313 \pm 0.05	0.5744 \pm 0.15	0.6351 \pm 0.08	22.6874 \pm 0.97	1.3428 \pm 0.65	2.5122 \pm 34.87
	$\lambda = 100.0$	15	1×10^{-5}	200	0.3477 \pm 0.03	0.2315 \pm 0.05	0.5770 \pm 0.15	0.6396 \pm 0.08	22.6840 \pm 0.98	1.3551 \pm 0.64	1.9321 \pm 31.62

H.5. Evaluation Results for Fine-Tuning: FLUX.1 [dev]

Table 17. **Part I:** Fine-tuning results for each heuristic on FLUX.1 [dev] across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, $Lambda$ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	N	Lr.	Ckpt.	CLIP I-T\uparrow	SigLIP-2 I-T\uparrow	BLIP T-T\uparrow	Qwen2 T-T\uparrow	PickScore I-T\uparrow	ImgRew I-T\uparrow	Composite\uparrow
	$\lambda = 0$	—	—	—	0.3449 \pm 0.03	0.2271 \pm 0.05	0.5739 \pm 0.15	0.6300 \pm 0.09	23.4234 \pm 1.03	1.2970 \pm 0.66	0.0000 \pm 0.00
Attend&Excite	$\lambda = 0.1$	1	5×10^{-5}	200	0.3463 \pm 0.03	0.2309 \pm 0.04	0.5804 \pm 0.15	0.6337 \pm 0.09	23.4598 \pm 1.02	1.3812 \pm 0.62	0.8987 \pm 78.81
	$\lambda = 0.1$	1	5×10^{-5}	250	0.3463 \pm 0.03	0.2304 \pm 0.05	0.5776 \pm 0.15	0.6343 \pm 0.09	23.4613 \pm 1.02	1.3794 \pm 0.63	0.7197 \pm 77.92
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3461 \pm 0.03	0.2320 \pm 0.05	0.5825 \pm 0.15	0.6363 \pm 0.08	23.3879 \pm 1.01	1.3826 \pm 0.63	2.0527 \pm 78.91
	$\lambda = 1.0$	1	5×10^{-5}	250	0.3460 \pm 0.03	0.2317 \pm 0.05	0.5826 \pm 0.14	0.6352 \pm 0.08	23.3893 \pm 1.00	1.3788 \pm 0.63	1.1855 \pm 76.88
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3468 \pm 0.03	0.2320 \pm 0.05	0.5876 \pm 0.15	0.6382 \pm 0.08	23.3333 \pm 1.01	1.3806 \pm 0.62	2.3477 \pm 79.28
	$\lambda = 10.0$	1	5×10^{-5}	250	0.3465 \pm 0.03	0.2324 \pm 0.05	0.5900 \pm 0.14	0.6358 \pm 0.08	23.3205 \pm 1.00	1.3696 \pm 0.64	2.0806 \pm 77.37
CONFORM	$\lambda = 0.1$	1	5×10^{-5}	200	0.3463 \pm 0.03	0.2295 \pm 0.05	0.5814 \pm 0.15	0.6369 \pm 0.09	23.4485 \pm 1.02	1.3530 \pm 0.63	1.3362 \pm 77.93
	$\lambda = 0.1$	1	5×10^{-5}	250	0.3464 \pm 0.03	0.2291 \pm 0.05	0.5738 \pm 0.15	0.6360 \pm 0.08	23.4443 \pm 1.03	1.3516 \pm 0.63	0.4789 \pm 75.25
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3463 \pm 0.03	0.2309 \pm 0.04	0.5831 \pm 0.15	0.6346 \pm 0.08	23.3804 \pm 0.99	1.3654 \pm 0.63	1.2216 \pm 81.01
	$\lambda = 1.0$	1	5×10^{-5}	250	0.3464 \pm 0.03	0.2308 \pm 0.04	0.5810 \pm 0.15	0.6328 \pm 0.09	23.3807 \pm 1.00	1.3627 \pm 0.63	1.0111 \pm 80.59
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3458 \pm 0.03	0.2305 \pm 0.04	0.5800 \pm 0.15	0.6369 \pm 0.08	23.3724 \pm 1.00	1.3631 \pm 0.63	1.9591 \pm 82.88
	$\lambda = 10.0$	1	5×10^{-5}	250	0.3460 \pm 0.03	0.2307 \pm 0.04	0.5833 \pm 0.15	0.6311 \pm 0.08	23.3696 \pm 1.00	1.3613 \pm 0.63	1.8922 \pm 84.73
Divide&Bind	$\lambda = 0.1$	1	5×10^{-5}	200	0.3445 \pm 0.03	0.2296 \pm 0.05	0.5705 \pm 0.15	0.6246 \pm 0.09	23.1909 \pm 1.06	1.2269 \pm 0.70	0.2002 \pm 47.34
	$\lambda = 0.1$	1	5×10^{-5}	250	0.3398 \pm 0.03	0.2254 \pm 0.05	0.5630 \pm 0.14	0.6127 \pm 0.09	22.9327 \pm 1.10	1.0971 \pm 0.76	-0.5948 \pm 47.17
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3414 \pm 0.03	0.2298 \pm 0.05	0.5718 \pm 0.14	0.6163 \pm 0.09	22.9084 \pm 1.09	1.1595 \pm 0.75	-0.5107 \pm 50.32
	$\lambda = 1.0$	1	5×10^{-5}	250	0.3400 \pm 0.03	0.2280 \pm 0.05	0.5674 \pm 0.15	0.6158 \pm 0.09	22.8820 \pm 1.07	1.1088 \pm 0.78	-1.7526 \pm 47.28
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3395 \pm 0.03	0.2289 \pm 0.05	0.5680 \pm 0.15	0.6111 \pm 0.09	22.6042 \pm 1.13	1.0909 \pm 0.79	-2.5737 \pm 47.12
	$\lambda = 10.0$	1	5×10^{-5}	250	0.3371 \pm 0.03	0.2238 \pm 0.05	0.5639 \pm 0.15	0.6046 \pm 0.09	22.5095 \pm 1.14	0.9932 \pm 0.82	-4.5041 \pm 52.16
Self-Cross Guidance	$\lambda = 0.1$	1	5×10^{-5}	200	0.3430 \pm 0.03	0.2259 \pm 0.05	0.5890 \pm 0.14	0.6169 \pm 0.09	22.9809 \pm 1.05	1.2671 \pm 0.71	-0.5809 \pm 59.68
	$\lambda = 0.1$	1	5×10^{-5}	250	0.3358 \pm 0.03	0.2172 \pm 0.04	0.5684 \pm 0.14	0.5938 \pm 0.09	22.4066 \pm 1.06	1.0405 \pm 0.81	-7.8207 \pm 65.80
	$\lambda = 0.1$	15	5×10^{-5}	200	0.3341 \pm 0.03	0.2022 \pm 0.05	0.5298 \pm 0.15	0.5743 \pm 0.10	22.0493 \pm 1.20	0.6552 \pm 1.02	-21.2190 \pm 100.86
	$\lambda = 0.1$	15	5×10^{-5}	250	0.3005 \pm 0.04	0.1439 \pm 0.06	0.4460 \pm 0.16	0.4704 \pm 0.13	20.3189 \pm 1.39	-0.8012 \pm 1.06	-48.7739 \pm 91.20
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3245 \pm 0.03	0.1926 \pm 0.05	0.5203 \pm 0.15	0.5491 \pm 0.12	20.9905 \pm 1.28	0.4266 \pm 1.08	-21.2172 \pm 79.42
	$\lambda = 1.0$	1	5×10^{-5}	250	0.2786 \pm 0.04	0.1108 \pm 0.05	0.3638 \pm 0.18	0.3842 \pm 0.15	19.0722 \pm 1.28	-1.4410 \pm 0.95	-67.6915 \pm 148.74
	$\lambda = 1.0$	15	5×10^{-5}	200	0.2954 \pm 0.04	0.1292 \pm 0.06	0.3778 \pm 0.16	0.4220 \pm 0.14	19.4072 \pm 1.27	-1.0540 \pm 1.02	-57.5407 \pm 134.08
	$\lambda = 1.0$	15	5×10^{-5}	250	0.2445 \pm 0.04	0.0730 \pm 0.05	0.2153 \pm 0.15	0.2824 \pm 0.10	18.3242 \pm 1.08	-2.0518 \pm 0.37	-81.8277 \pm 175.76
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3212 \pm 0.03	0.1859 \pm 0.05	0.5159 \pm 0.16	0.5429 \pm 0.12	20.8077 \pm 1.29	0.3106 \pm 1.11	-28.5637 \pm 104.51
	$\lambda = 10.0$	1	5×10^{-5}	250	0.2627 \pm 0.04	0.0925 \pm 0.05	0.3145 \pm 0.18	0.3391 \pm 0.15	18.6899 \pm 1.24	-1.7644 \pm 0.79	-74.5966 \pm 159.72
	$\lambda = 10.0$	15	5×10^{-5}	200	0.2878 \pm 0.04	0.1166 \pm 0.05	0.3505 \pm 0.17	0.3889 \pm 0.14	18.9787 \pm 1.17	-1.3775 \pm 0.91	-65.3953 \pm 126.10
	$\lambda = 10.0$	15	5×10^{-5}	250	0.2363 \pm 0.03	0.0583 \pm 0.04	0.1816 \pm 0.13	0.2557 \pm 0.10	18.0894 \pm 1.09	-2.1576 \pm 0.26	-85.8370 \pm 187.00
FOCUS(Ours)	$\lambda = 0.1$	1	5×10^{-5}	200	0.3470 \pm 0.03	0.2315 \pm 0.05	0.5730 \pm 0.15	0.6354 \pm 0.08	23.4339 \pm 1.03	1.3893 \pm 0.62	0.7730 \pm 79.23
	$\lambda = 0.1$	1	5×10^{-5}	250	0.3465 \pm 0.03	0.2309 \pm 0.04	0.5764 \pm 0.15	0.6331 \pm 0.09	23.4301 \pm 1.02	1.3826 \pm 0.63	1.7561 \pm 80.48
	$\lambda = 0.1$	1	5×10^{-5}	300	0.3467 \pm 0.03	0.2311 \pm 0.05	0.5746 \pm 0.15	0.6364 \pm 0.08	23.4418 \pm 1.02	1.3837 \pm 0.63	1.2623 \pm 79.78
	$\lambda = 0.1$	15	5×10^{-5}	200	0.3478 \pm 0.03	0.2323 \pm 0.05	0.5751 \pm 0.15	0.6271 \pm 0.08	23.4197 \pm 1.03	1.3501 \pm 0.66	0.4181 \pm 86.29
	$\lambda = 0.1$	15	5×10^{-5}	250	0.3477 \pm 0.03	0.2317 \pm 0.05	0.5774 \pm 0.14	0.6299 \pm 0.08	23.4069 \pm 1.01	1.3398 \pm 0.67	0.2517 \pm 85.36
	$\lambda = 0.1$	15	5×10^{-5}	300	0.3477 \pm 0.03	0.2310 \pm 0.05	0.5800 \pm 0.14	0.6288 \pm 0.09	23.3904 \pm 1.01	1.3215 \pm 0.69	-0.2052 \pm 89.69
	$\lambda = 0.1$	150	5×10^{-5}	200	0.3483 \pm 0.03	0.2337 \pm 0.05	0.5776 \pm 0.14	0.6313 \pm 0.09	23.3797 \pm 1.03	1.3696 \pm 0.65	1.8100 \pm 90.13
	$\lambda = 0.1$	150	5×10^{-5}	250	0.3472 \pm 0.03	0.2329 \pm 0.05	0.5766 \pm 0.15	0.6249 \pm 0.09	23.3419 \pm 1.03	1.3375 \pm 0.67	0.0200 \pm 91.30
	$\lambda = 0.1$	1	1×10^{-4}	200	0.3463 \pm 0.03	0.2306 \pm 0.05	0.5716 \pm 0.15	0.6352 \pm 0.09	23.4556 \pm 1.02	1.3858 \pm 0.63	1.6167 \pm 82.35
	$\lambda = 0.1$	1	1×10^{-4}	250	0.3460 \pm 0.03	0.2303 \pm 0.05	0.5706 \pm 0.15	0.6373 \pm 0.08	23.4515 \pm 1.02	1.3804 \pm 0.63	1.4079 \pm 85.07
	$\lambda = 0.1$	1	1×10^{-4}	300	0.3461 \pm 0.03	0.2301 \pm 0.05	0.5712 \pm 0.15	0.6351 \pm 0.08	23.4514 \pm 1.02	1.3693 \pm 0.64	0.5386 \pm 83.39
	$\lambda = 0.1$	15	1×10^{-4}	200	0.3466 \pm 0.03	0.2290 \pm 0.05	0.5791 \pm 0.14	0.6252 \pm 0.09	23.3434 \pm 1.01	1.2905 \pm 0.72	-0.5905 \pm 78.21
	$\lambda = 0.1$	15	1×10^{-4}	250	0.3458 \pm 0.03	0.2278 \pm 0.05	0.5754 \pm 0.14	0.6255 \pm 0.09	23.3281 \pm 1.02	1.2646 \pm 0.73	-1.0812 \pm 76.48
	$\lambda = 0.1$	15	1×10^{-4}	300	0.3461 \pm 0.03	0.2275 \pm 0.05	0.5763 \pm 0.14	0.6245 \pm 0.09	23.3188 \pm 1.01	1.2586 \pm 0.73	-0.8369 \pm 76.50
	$\lambda = 0.1$	1	1×10^{-5}	200	0.3447 \pm 0.03	0.2274 \pm 0.05	0.5699 \pm 0.15	0.6319 \pm 0.09	23.4495 \pm 1.03	1.3103 \pm 0.66	0.6586 \pm 17.62
	$\lambda = 0.1$	1	1×10^{-5}	250	0.3451 \pm 0.03	0.2280 \pm 0.05	0.5716 \pm 0.15	0.6307 \pm 0.09	23.4529 \pm 1.02	1.3200 \pm 0.65	1.1095 \pm 20.82
	$\lambda = 0.1$	1	1×10^{-5}	300	0.3455 \pm 0.03	0.2287 \pm 0.05	0.5721 \pm 0.16	0.6324 \pm 0.08	23.4562 \pm 1.03	1.3348 \pm 0.63	1.4195 \pm 28.92
	$\lambda = 0.1$	15	1×10^{-5}	200	0.3449 \pm 0.03	0.2275 \pm 0.05	0.5715 \pm 0.15	0.6344 \pm 0.08	23.4299 \pm 1.02	1.3063 \pm 0.65	0.7188 \pm 12.34
	$\lambda = 0.1$	15	1×10^{-5}	250	0.3448 \pm 0.03	0.2274 \pm 0.05	0.5720 \pm 0.15	0.6310 \pm 0.08	23.4378 \pm 1.03	1.3064 \pm 0.66	0.7701 \pm 14.62

Table 18. **Part II:** Fine-tuning results for each heuristic on FLUX.1 [dev] across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, $Lambda$ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	N	Lr.	Ckpt.	CLIP I-T\uparrow	SigLIP-2 I-T\uparrow	BLIP T-T\uparrow	Qwen2 T-T\uparrow	PickScore I-T\uparrow	ImgRew I-T\uparrow	Composite\uparrow
FOCUS(Ours)	$\lambda = 0.1$	15	1×10^{-5}	300	0.3448 \pm 0.03	0.2276 \pm 0.05	0.5741 \pm 0.15	0.6320 \pm 0.09	23.4507 \pm 1.03	1.3135 \pm 0.65	1.3703 \pm 24.39
	$\lambda = 1.0$	1	5×10^{-5}	200	0.3468 \pm 0.03	0.2321 \pm 0.04	0.5802 \pm 0.15	0.6351 \pm 0.09	23.3666 \pm 1.00	1.3870 \pm 0.62	1.7951 \pm 78.85
	$\lambda = 1.0$	1	5×10^{-5}	250	0.3467 \pm 0.03	0.2324 \pm 0.05	0.5807 \pm 0.15	0.6335 \pm 0.08	23.3580 \pm 1.01	1.3937 \pm 0.61	2.2221 \pm 80.05
	$\lambda = 1.0$	1	5×10^{-5}	300	0.3468 \pm 0.03	0.2321 \pm 0.05	0.5810 \pm 0.15	0.6335 \pm 0.08	23.3571 \pm 1.01	1.3858 \pm 0.62	2.1067 \pm 80.98
	$\lambda = 1.0$	15	5×10^{-5}	200	0.3468 \pm 0.03	0.2302 \pm 0.05	0.5781 \pm 0.14	0.6207 \pm 0.09	23.2315 \pm 1.02	1.2788 \pm 0.71	-1.5300 \pm 85.78
	$\lambda = 1.0$	15	5×10^{-5}	250	0.3447 \pm 0.03	0.2273 \pm 0.05	0.5638 \pm 0.15	0.6162 \pm 0.09	23.1107 \pm 1.06	1.1971 \pm 0.76	-5.9229 \pm 77.12
	$\lambda = 1.0$	15	5×10^{-5}	300	0.3422 \pm 0.03	0.2207 \pm 0.05	0.5595 \pm 0.14	0.6039 \pm 0.09	22.9404 \pm 1.10	1.0819 \pm 0.82	-8.3086 \pm 77.13
	$\lambda = 1.0$	150	5×10^{-5}	200	0.3469 \pm 0.03	0.2320 \pm 0.05	0.5799 \pm 0.15	0.6226 \pm 0.09	23.1509 \pm 1.03	1.2994 \pm 0.71	-1.4642 \pm 85.91
	$\lambda = 1.0$	150	5×10^{-5}	250	0.3385 \pm 0.03	0.2168 \pm 0.05	0.5544 \pm 0.14	0.5906 \pm 0.10	22.4382 \pm 1.12	0.9719 \pm 0.84	-10.4580 \pm 80.37
	$\lambda = 1.0$	1	1×10^{-4}	200	0.3459 \pm 0.03	0.2312 \pm 0.05	0.5792 \pm 0.15	0.6333 \pm 0.09	23.3473 \pm 1.04	1.3823 \pm 0.62	1.1661 \pm 81.85
	$\lambda = 1.0$	1	1×10^{-4}	250	0.3454 \pm 0.03	0.2308 \pm 0.05	0.5783 \pm 0.15	0.6358 \pm 0.08	23.3585 \pm 1.03	1.3864 \pm 0.62	2.0272 \pm 80.80
	$\lambda = 1.0$	1	1×10^{-4}	300	0.3453 \pm 0.03	0.2309 \pm 0.05	0.5790 \pm 0.15	0.6348 \pm 0.08	23.3754 \pm 1.03	1.3872 \pm 0.63	1.7788 \pm 80.38
	$\lambda = 1.0$	15	1×10^{-4}	200	0.3300 \pm 0.03	0.1979 \pm 0.06	0.5211 \pm 0.15	0.5701 \pm 0.11	22.3704 \pm 1.26	0.5668 \pm 1.01	-18.8541 \pm 77.72
	$\lambda = 1.0$	15	1×10^{-4}	250	0.3234 \pm 0.03	0.1859 \pm 0.06	0.5010 \pm 0.15	0.5482 \pm 0.11	22.0612 \pm 1.34	0.2713 \pm 1.05	-23.5034 \pm 78.95
	$\lambda = 1.0$	15	1×10^{-4}	300	0.3228 \pm 0.03	0.1835 \pm 0.05	0.5049 \pm 0.15	0.5434 \pm 0.11	22.0043 \pm 1.30	0.2106 \pm 1.06	-25.7103 \pm 78.11
	$\lambda = 1.0$	1	1×10^{-5}	200	0.3447 \pm 0.03	0.2276 \pm 0.05	0.5757 \pm 0.15	0.6329 \pm 0.08	23.4457 \pm 1.03	1.3155 \pm 0.64	0.5969 \pm 19.40
	$\lambda = 1.0$	1	1×10^{-5}	250	0.3452 \pm 0.03	0.2281 \pm 0.05	0.5735 \pm 0.15	0.6294 \pm 0.08	23.4471 \pm 1.03	1.3236 \pm 0.64	1.7384 \pm 28.80
	$\lambda = 1.0$	1	1×10^{-5}	300	0.3454 \pm 0.03	0.2289 \pm 0.05	0.5728 \pm 0.15	0.6329 \pm 0.08	23.4613 \pm 1.03	1.3439 \pm 0.63	-0.2485 \pm 68.61
	$\lambda = 1.0$	15	1×10^{-5}	200	0.3449 \pm 0.03	0.2273 \pm 0.05	0.5686 \pm 0.15	0.6315 \pm 0.09	23.4335 \pm 1.03	1.3003 \pm 0.66	0.3974 \pm 11.86
	$\lambda = 1.0$	15	1×10^{-5}	250	0.3450 \pm 0.03	0.2275 \pm 0.05	0.5705 \pm 0.15	0.6316 \pm 0.08	23.4428 \pm 1.04	1.3082 \pm 0.66	1.1473 \pm 15.58
	$\lambda = 1.0$	15	1×10^{-5}	300	0.3449 \pm 0.03	0.2277 \pm 0.05	0.5730 \pm 0.15	0.6331 \pm 0.09	23.4532 \pm 1.03	1.3132 \pm 0.65	1.2816 \pm 23.44
	$\lambda = 10.0$	1	5×10^{-5}	200	0.3464 \pm 0.03	0.2324 \pm 0.04	0.5793 \pm 0.15	0.6343 \pm 0.08	23.3726 \pm 1.00	1.3885 \pm 0.61	0.7007 \pm 80.04
	$\lambda = 10.0$	1	5×10^{-5}	250	0.3465 \pm 0.03	0.2322 \pm 0.05	0.5830 \pm 0.15	0.6343 \pm 0.08	23.3351 \pm 1.02	1.3845 \pm 0.62	1.8589 \pm 76.17
	$\lambda = 10.0$	1	5×10^{-5}	300	0.3465 \pm 0.03	0.2320 \pm 0.04	0.5795 \pm 0.15	0.6343 \pm 0.08	23.3228 \pm 1.01	1.3772 \pm 0.62	1.4944 \pm 77.16
	$\lambda = 10.0$	15	5×10^{-5}	200	0.3461 \pm 0.03	0.2297 \pm 0.05	0.5724 \pm 0.15	0.6187 \pm 0.09	23.2078 \pm 1.02	1.2722 \pm 0.71	-1.6496 \pm 87.09
	$\lambda = 10.0$	15	5×10^{-5}	250	0.3447 \pm 0.03	0.2265 \pm 0.05	0.5649 \pm 0.15	0.6145 \pm 0.09	23.1177 \pm 1.06	1.2000 \pm 0.78	-6.2880 \pm 81.26
	$\lambda = 10.0$	15	5×10^{-5}	300	0.3402 \pm 0.03	0.2189 \pm 0.05	0.5551 \pm 0.15	0.6018 \pm 0.10	22.8512 \pm 1.12	1.0158 \pm 0.82	-10.0126 \pm 77.64
	$\lambda = 10.0$	150	5×10^{-5}	200	0.3421 \pm 0.03	0.2277 \pm 0.05	0.5610 \pm 0.15	0.6082 \pm 0.10	22.7521 \pm 1.04	1.1727 \pm 0.76	-4.4918 \pm 82.06
	$\lambda = 10.0$	150	5×10^{-5}	250	0.3288 \pm 0.03	0.1920 \pm 0.05	0.5076 \pm 0.16	0.5499 \pm 0.12	21.5018 \pm 1.26	0.4431 \pm 1.06	-24.8389 \pm 107.26
	$\lambda = 10.0$	1	1×10^{-4}	200	0.3451 \pm 0.03	0.2306 \pm 0.04	0.5819 \pm 0.15	0.6328 \pm 0.08	23.2743 \pm 1.01	1.3617 \pm 0.65	1.9009 \pm 81.53
	$\lambda = 10.0$	1	1×10^{-4}	250	0.3449 \pm 0.03	0.2302 \pm 0.04	0.5857 \pm 0.15	0.6310 \pm 0.08	23.2640 \pm 1.01	1.3596 \pm 0.65	2.0116 \pm 80.29
	$\lambda = 10.0$	1	1×10^{-4}	300	0.3447 \pm 0.03	0.2298 \pm 0.04	0.5784 \pm 0.15	0.6299 \pm 0.08	23.2576 \pm 1.01	1.3588 \pm 0.65	1.2325 \pm 80.15
	$\lambda = 10.0$	15	1×10^{-4}	200	0.3211 \pm 0.04	0.1712 \pm 0.05	0.4878 \pm 0.15	0.5311 \pm 0.12	21.5464 \pm 1.38	0.0696 \pm 1.09	-29.9341 \pm 78.90
	$\lambda = 10.0$	15	1×10^{-4}	250	0.3050 \pm 0.04	0.1496 \pm 0.05	0.4513 \pm 0.16	0.4905 \pm 0.12	20.7941 \pm 1.44	-0.4887 \pm 1.14	-41.0344 \pm 81.71
	$\lambda = 10.0$	15	1×10^{-4}	300	0.3016 \pm 0.04	0.1443 \pm 0.06	0.4421 \pm 0.16	0.4796 \pm 0.13	20.7355 \pm 1.42	-0.5999 \pm 1.12	-43.8700 \pm 88.32
	$\lambda = 10.0$	1	1×10^{-5}	200	0.3448 \pm 0.03	0.2276 \pm 0.05	0.5735 \pm 0.15	0.6313 \pm 0.08	23.4506 \pm 1.03	1.3128 \pm 0.65	0.6894 \pm 19.58
	$\lambda = 10.0$	1	1×10^{-5}	250	0.3453 \pm 0.03	0.2280 \pm 0.05	0.5721 \pm 0.15	0.6329 \pm 0.08	23.4550 \pm 1.02	1.3244 \pm 0.64	1.9160 \pm 23.65
	$\lambda = 10.0$	1	1×10^{-5}	300	0.3453 \pm 0.03	0.2288 \pm 0.05	0.5738 \pm 0.15	0.6342 \pm 0.09	23.4675 \pm 1.03	1.3432 \pm 0.63	0.2377 \pm 61.73
	$\lambda = 10.0$	15	1×10^{-5}	200	0.3449 \pm 0.03	0.2274 \pm 0.05	0.5713 \pm 0.15	0.6324 \pm 0.08	23.4273 \pm 1.02	1.2991 \pm 0.66	0.2980 \pm 10.93
	$\lambda = 10.0$	15	1×10^{-5}	250	0.3449 \pm 0.03	0.2277 \pm 0.05	0.5698 \pm 0.15	0.6354 \pm 0.08	23.4442 \pm 1.03	1.3056 \pm 0.66	1.1876 \pm 18.26
$\lambda = 10.0$	15	1×10^{-5}	300	0.3449 \pm 0.03	0.2276 \pm 0.05	0.5673 \pm 0.15	0.6323 \pm 0.09	23.4523 \pm 1.04	1.3127 \pm 0.65	0.8921 \pm 23.74	
$\lambda = 0.01$	1	5×10^{-5}	200	0.3457 \pm 0.03	0.2298 \pm 0.04	0.5737 \pm 0.15	0.6336 \pm 0.08	23.4510 \pm 1.03	1.3584 \pm 0.64	0.2960 \pm 80.75	
$\lambda = 0.01$	1	5×10^{-5}	250	0.3462 \pm 0.03	0.2297 \pm 0.05	0.5741 \pm 0.15	0.6348 \pm 0.09	23.4553 \pm 1.02	1.3605 \pm 0.64	-0.2550 \pm 74.33	
$\lambda = 0.01$	1	5×10^{-5}	300	0.3459 \pm 0.03	0.2299 \pm 0.05	0.5753 \pm 0.15	0.6366 \pm 0.09	23.4489 \pm 1.02	1.3656 \pm 0.64	-0.1884 \pm 77.38	
$\lambda = 0.01$	15	5×10^{-5}	200	0.3470 \pm 0.03	0.2322 \pm 0.05	0.5826 \pm 0.14	0.6330 \pm 0.09	23.4871 \pm 1.02	1.3995 \pm 0.62	2.4565 \pm 86.64	
$\lambda = 0.01$	15	5×10^{-5}	250	0.3466 \pm 0.03	0.2310 \pm 0.05	0.5800 \pm 0.14	0.6281 \pm 0.09	23.4746 \pm 1.00	1.3663 \pm 0.64	1.5619 \pm 80.91	
$\lambda = 0.01$	15	5×10^{-5}	300	0.3466 \pm 0.03	0.2309 \pm 0.05	0.5771 \pm 0.14	0.6313 \pm 0.09	23.4660 \pm 1.02	1.3493 \pm 0.66	1.4770 \pm 83.80	
$\lambda = 0.01$	1	1×10^{-4}	200	0.3458 \pm 0.03	0.2291 \pm 0.05	0.5774 \pm 0.15	0.6340 \pm 0.08	23.4498 \pm 1.01	1.3470 \pm 0.66	-0.4316 \pm 72.59	
$\lambda = 0.01$	1	1×10^{-4}	250	0.3454 \pm 0.03	0.2280 \pm 0.05	0.5729 \pm 0.15	0.6338 \pm 0.08	23.4481 \pm 1.01	1.3332 \pm 0.67	-0.7144 \pm 73.68	
$\lambda = 0.01$	1	1×10^{-4}	300	0.3454 \pm 0.03	0.2278 \pm 0.05	0.5715 \pm 0.15	0.6339 \pm 0.08	23.4431 \pm 1.03	1.3273 \pm 0.67	-0.4719 \pm 71.06	

Table 19. **Part III:** Fine-tuning results for each heuristic on FLUX.1 [dev] across different hyperparameter configurations. Here, N denotes the number of prompts in the dataset, Lr the learning rate, $Lambda$ the scalar applied to the heuristic function, and $Ckpt.$ the checkpoint used for evaluation. We report mean \pm std over all prompts and seeds. All methods use the same sampling and evaluation pipeline. Values are color-coded relative to the base case: entries close to the base value are white, improvements are shown in blue and degradations in red, with stronger color intensity indicating larger deviations from the base.

	Lambda	N	Lr.	Ckpt.	CLIP I-T\uparrow	SigLIP-2 I-T\uparrow	BLIP T-T\uparrow	Qwen2 T-T\uparrow	PickScore I-T\uparrow	ImgRew I-T\uparrow	Composite\uparrow
	$\lambda = 0.01$	15	1×10^{-4}	200	0.3458 \pm 0.03	0.2298 \pm 0.05	0.5734 \pm 0.15	0.6277 \pm 0.09	23.4402 \pm 1.02	1.3299 \pm 0.67	0.9600 \pm 82.71
	$\lambda = 0.01$	15	1×10^{-4}	250	0.3449 \pm 0.03	0.2276 \pm 0.05	0.5749 \pm 0.15	0.6280 \pm 0.09	23.4544 \pm 1.01	1.2758 \pm 0.71	-1.2966 \pm 78.75
	$\lambda = 0.01$	15	1×10^{-4}	300	0.3436 \pm 0.03	0.2260 \pm 0.05	0.5700 \pm 0.15	0.6257 \pm 0.09	23.4034 \pm 1.01	1.2353 \pm 0.74	-2.8836 \pm 80.46
	$\lambda = 0.01$	1	1×10^{-5}	200	0.3448 \pm 0.03	0.2275 \pm 0.05	0.5702 \pm 0.15	0.6322 \pm 0.08	23.4298 \pm 1.02	1.3062 \pm 0.65	0.2255 \pm 11.26
	$\lambda = 0.01$	1	1×10^{-5}	250	0.3448 \pm 0.03	0.2275 \pm 0.05	0.5695 \pm 0.15	0.6309 \pm 0.08	23.4399 \pm 1.03	1.3053 \pm 0.66	0.5886 \pm 13.64
	$\lambda = 0.01$	1	1×10^{-5}	300	0.3448 \pm 0.03	0.2276 \pm 0.05	0.5723 \pm 0.15	0.6300 \pm 0.08	23.4483 \pm 1.03	1.3139 \pm 0.65	0.6849 \pm 15.81
	$\lambda = 0.01$	15	1×10^{-5}	200	0.3450 \pm 0.03	0.2275 \pm 0.05	0.5708 \pm 0.15	0.6315 \pm 0.08	23.4302 \pm 1.02	1.3007 \pm 0.66	-0.1373 \pm 9.07
	$\lambda = 0.01$	15	1×10^{-5}	250	0.3452 \pm 0.03	0.2274 \pm 0.05	0.5706 \pm 0.15	0.6302 \pm 0.09	23.4373 \pm 1.03	1.3074 \pm 0.66	0.3545 \pm 9.44
	$\lambda = 0.01$	15	1×10^{-5}	300	0.3449 \pm 0.03	0.2275 \pm 0.05	0.5689 \pm 0.15	0.6315 \pm 0.08	23.4320 \pm 1.02	1.3073 \pm 0.66	0.3125 \pm 10.96
	$\lambda = 100.0$	1	5×10^{-5}	200	0.3467 \pm 0.03	0.2324 \pm 0.04	0.5831 \pm 0.15	0.6348 \pm 0.08	23.3717 \pm 1.00	1.3933 \pm 0.61	2.0896 \pm 83.17
	$\lambda = 100.0$	1	5×10^{-5}	250	0.3468 \pm 0.03	0.2328 \pm 0.05	0.5780 \pm 0.15	0.6386 \pm 0.08	23.3278 \pm 1.01	1.3899 \pm 0.61	2.5881 \pm 78.83
	$\lambda = 100.0$	1	5×10^{-5}	300	0.3467 \pm 0.03	0.2320 \pm 0.04	0.5807 \pm 0.15	0.6346 \pm 0.08	23.3001 \pm 1.00	1.3830 \pm 0.62	2.0565 \pm 75.73
	$\lambda = 100.0$	15	5×10^{-5}	200	0.3468 \pm 0.03	0.2329 \pm 0.05	0.5748 \pm 0.14	0.6203 \pm 0.09	23.1693 \pm 1.03	1.2929 \pm 0.71	-1.5988 \pm 86.77
	$\lambda = 100.0$	15	5×10^{-5}	250	0.3445 \pm 0.03	0.2295 \pm 0.05	0.5618 \pm 0.15	0.6151 \pm 0.09	23.0459 \pm 1.04	1.2096 \pm 0.76	-5.6673 \pm 81.37
	$\lambda = 100.0$	15	5×10^{-5}	300	0.3409 \pm 0.03	0.2223 \pm 0.05	0.5516 \pm 0.15	0.6008 \pm 0.10	22.7607 \pm 1.12	1.0481 \pm 0.83	-8.5041 \pm 79.04
	$\lambda = 100.0$	150	5×10^{-5}	200	0.3452 \pm 0.03	0.2334 \pm 0.05	0.5757 \pm 0.15	0.6190 \pm 0.09	23.0172 \pm 1.01	1.2729 \pm 0.72	-2.0022 \pm 88.56
	$\lambda = 100.0$	150	5×10^{-5}	250	0.3347 \pm 0.03	0.2127 \pm 0.05	0.5420 \pm 0.15	0.5843 \pm 0.11	22.1857 \pm 1.16	0.8392 \pm 0.93	-13.6503 \pm 79.56
	$\lambda = 100.0$	1	1×10^{-4}	200	0.3452 \pm 0.03	0.2300 \pm 0.04	0.5787 \pm 0.15	0.6305 \pm 0.08	23.2591 \pm 1.01	1.3748 \pm 0.65	1.9560 \pm 84.13
	$\lambda = 100.0$	1	1×10^{-4}	250	0.3456 \pm 0.03	0.2294 \pm 0.04	0.5810 \pm 0.15	0.6311 \pm 0.08	23.2596 \pm 1.00	1.3674 \pm 0.64	1.8451 \pm 83.33
	$\lambda = 100.0$	1	1×10^{-4}	300	0.3452 \pm 0.03	0.2294 \pm 0.04	0.5794 \pm 0.15	0.6256 \pm 0.09	23.2467 \pm 1.00	1.3590 \pm 0.65	1.1347 \pm 81.41
	$\lambda = 100.0$	15	1×10^{-4}	200	0.2524 \pm 0.04	0.0866 \pm 0.05	0.2469 \pm 0.15	0.3093 \pm 0.14	18.5154 \pm 1.23	-1.8946 \pm 0.67	-76.1686 \pm 190.16
	$\lambda = 100.0$	15	1×10^{-4}	250	0.2397 \pm 0.04	0.0773 \pm 0.05	0.2278 \pm 0.14	0.2825 \pm 0.12	18.1672 \pm 1.17	-2.0543 \pm 0.49	-81.0538 \pm 192.97
	$\lambda = 100.0$	15	1×10^{-4}	300	0.2338 \pm 0.04	0.0720 \pm 0.05	0.2232 \pm 0.13	0.2685 \pm 0.11	17.9936 \pm 1.14	-2.1224 \pm 0.38	-83.1029 \pm 193.91
	$\lambda = 100.0$	1	1×10^{-5}	200	0.3446 \pm 0.03	0.2277 \pm 0.05	0.5729 \pm 0.15	0.6322 \pm 0.08	23.4473 \pm 1.03	1.3107 \pm 0.65	0.7338 \pm 18.91
	$\lambda = 100.0$	1	1×10^{-5}	250	0.3451 \pm 0.03	0.2281 \pm 0.05	0.5742 \pm 0.15	0.6340 \pm 0.08	23.4579 \pm 1.02	1.3294 \pm 0.64	1.4770 \pm 22.37
	$\lambda = 100.0$	1	1×10^{-5}	300	0.3455 \pm 0.03	0.2287 \pm 0.05	0.5741 \pm 0.15	0.6322 \pm 0.08	23.4750 \pm 1.03	1.3416 \pm 0.63	-0.9838 \pm 68.16
	$\lambda = 100.0$	15	1×10^{-5}	200	0.3447 \pm 0.03	0.2274 \pm 0.05	0.5709 \pm 0.15	0.6298 \pm 0.09	23.4367 \pm 1.02	1.3020 \pm 0.66	0.4266 \pm 10.89
	$\lambda = 100.0$	15	1×10^{-5}	250	0.3450 \pm 0.03	0.2276 \pm 0.05	0.5690 \pm 0.15	0.6333 \pm 0.09	23.4444 \pm 1.03	1.3102 \pm 0.65	0.9446 \pm 16.17
	$\lambda = 100.0$	15	1×10^{-5}	300	0.3451 \pm 0.03	0.2276 \pm 0.05	0.5699 \pm 0.15	0.6319 \pm 0.08	23.4522 \pm 1.03	1.3107 \pm 0.65	0.8363 \pm 20.47

FOCUS(Ours)

