CLOSER: CLOSING THE FEEDBACK LOOP FOR INFERENCE-TIME SCALING OF DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Inference-time scaling aims to improve diffusion-model outputs by allowing more time and computation during generation. Using task feedback to direct this computation requires turning a terminal assessment into an effective correction. Our controlled diagnostics show the challenge: informative evaluator gradients can still yield harmful updates, and the effective intervention stage varies across objectives. We therefore introduce **CLOSER**, a training-free controller that turns terminal feedback into verified, position-aware state corrections. At an initial or intermediate sampler state, CLOSER differentiates through the remaining discrete sampler to construct a bounded candidate set. It executes these candidates through the unchanged suffix and accepts a correction only after terminal evaluation, keeping the incoming trajectory available for rollback. After acceptance, refreshed feedback guides the next proposal from the updated sampler state. This separation lets gradients guide the search while terminal criteria determine which trajectory is retained. Across SD3.5-Medium and Z-Image on DPG-Bench, OneIG-EN, and GenEval, CLOSER achieves the strongest reported terminal utility across all evaluated backbone–benchmark settings, exceeding the strongest GenEval baselines by 0.113 and 0.106. The evaluated operating points also outperform the compared utility–cost baselines. Position analysis connects these gains to correction scale: early interventions favor structural changes, whereas late interventions favor preference and alignment refinement in the controlled study.

1 INTRODUCTION

Inference-time scaling seeks better outputs from frozen diffusion models by spending additional computation during generation. How effectively this budget is used depends on the signals that guide the sampling trajectory. Generator-native methods use the pretrained conditional field through classifier-free guidance, denoising–inversion reflection, or weak-to-strong differences (Ho & Salimans, 2022; Bai et al., 2025; 2026). Evaluator-access search introduces task-specific feedback by scoring seeds, partial trajectories, or completed images (Ma et al., 2025; Guimaraes & Perona, 2026). Evaluator-gradient methods use such feedback more directly, differentiating through generation to optimize initial noise or guide sampler states (Bansal et al., 2023; Yu et al., 2023; Ye et al., 2024; Guo et al., 2024; Wallace et al., 2023; Eyring et al., 2024). These approaches provide ways to allocate computation and propose changes. To turn a proposal into a reliable correction, however, its completed outcome must also determine whether the state is retained and how the next proposal is constructed.

We examine this control problem with a pilot that varies the correction signal, terminal verification, and intervention position. Over the tested compute range, generator-native and endpoint-search baselines saturate or become non-monotone, whereas terminally verified state interventions continue to improve utility (Figure 1(**A**)). The remaining diagnostics clarify this contrast. An informative proposal need not produce a beneficial update: direct execution can reduce terminal utility, while suffix evaluation and rollback prevent harmful accepted corrections (Figure 1(**B**)). Verification nevertheless benefits from an informed search direction, as evaluator-derived proposals outperform matched-random controls (Figure 1(**C**)). The final diagnostic asks where to apply that direction: in the SD3.5M study, early intervention favors structure, while late intervention favors preference and alignment (Figure 1(**D**)).

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

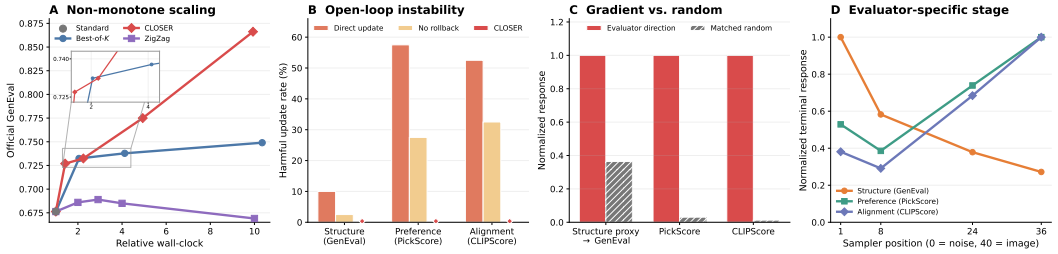


Figure 1: **The diagnostic path to closed-loop, stage-aware control. (A)** Over the tested compute range, generator-native and endpoint-search baselines saturate or become non-monotone, while verified state search continues to improve utility. **(B)** Direct evaluator-gradient updates can reduce terminal utility; suffix evaluation and rollback prevent harmful accepted updates. **(C)** Evaluator-derived directions outperform matched-random directions, separating direction information from candidate multiplicity. **(D)** The SD3.5M position study shows why placement also matters: early intervention favors structural correction, while late intervention favors preference and alignment refinement.

Together, these findings identify two requirements for using extra inference compute: verify what a proposed correction achieves, and apply it at a position suited to the objective. We study these requirements under **evaluator-access inference-time scaling**, where a frozen task evaluator can be queried during generation and defines the operational terminal utility. The goal is to convert this terminal assessment into control of the current trajectory: propose a change, execute its remaining sampling steps, and use the completed outcome to decide whether to continue from the corrected state.

We introduce **CLOSER**, an evaluator-grounded controller that implements this cycle at initial or intermediate sampler states. It differentiates a proposal objective through the remaining discrete sampler to construct a bounded set of local interventions. The unchanged suffix then realizes each candidate, and a terminal selector with an optional preservation guard chooses a feasible improvement or retains the incoming state. An accepted correction becomes the starting point for the next proposal, whose gradient is recomputed on the updated trajectory. Thus, *gradients propose; completed rollouts decide*. To support this division of labor, the framework separates proposal objective J, terminal selector V, preservation guard H, and reporting metric M. This permits differentiable task feedback and terminal preservation criteria to be combined without retraining any model.

Our contributions are:

- We formulate evaluator-access generation as closed-loop control over transient sampler states and introduce CLOSER, which links proposal, sampler realization, terminal acceptance, and rollback at both initial and intermediate positions.
- We develop a stage-aware exact-suffix intervention mechanism that propagates terminal feedback through the implemented sampler, constructs bounded local candidates, and refreshes feedback after accepted corrections. Position-dependent response analysis links intervention placement to the scale of the desired correction.
- Across two frozen generators and three benchmark families, CLOSER ranks first in all six settings, with its reported operating points above the compared utility–cost baselines. It improves the strongest GenEval baselines by $+0.113/+0.106$ and DPG-Bench baselines by $+7.38/+6.66$ points on SD3.5M/Z-Image. In the SD3.5M diagnostics, random-probe low-frequency reach contracts $10.43\times$ between early and late groups; early structural gain is $3.7\times$ larger, while late preference/alignment gains are $1.9\times/2.6\times$ larger.

2 Related Work

Inference-time scaling and feedback placement. Generator-native methods modify sampling with signals from the pretrained model: CFG amplifies conditional scores, while Zigzag and W2S reuse

self-reflective or weak-to-strong differences along the denoising path (Ho & Salimans, 2022; Bai et al., 2025; 2026). Evaluator-access search uses an external score to choose among alternatives. Best-of-K ranks completed samples, whereas progressive pruning allocates further denoising compute to promising partial trajectories (Ma et al., 2025; Guimaraes & Perona, 2026). CLOSER also uses terminal selection, but applies it to local corrections of a running state and continues from the accepted correction.

Evaluator-driven guidance and state optimization. To construct changes rather than select existing trajectories, training-free guidance injects predictor gradients into diffusion sampling (Bansal et al., 2023; Yu et al., 2023; Ye et al., 2024). InitNO, DOODL, and ReNO optimize initial latents using attention-derived or reward-based objectives (Guo et al., 2024; Wallace et al., 2023; Eyring et al., 2024). CLOSER connects this use of gradients with terminal verification at selected initial or intermediate positions. Each correction is evaluated through the remaining sampler, and its acceptance determines the state from which the next gradient is computed. The distinction is therefore the coupling of state search and feedback refresh, not gradient access alone.

Task evaluators and preservation criteria. This separation of proposal and acceptance allows different evaluators to serve different purposes. CLIPScore, PickScore, HPSv2, and ImageReward provide alignment or preference signals (Hessel et al., 2021; Kirstain et al., 2023; Wu et al., 2023; Xu et al., 2023), while GenEval, DPG-Bench, and OneIG-Bench measure structured prompt following and composition (Ghosh et al., 2023; Hu et al., 2024; Chang et al., 2025). CLOSER assigns these signals to proposal, selection, preservation, and reporting roles, so optimizing a task score can be combined with explicit auxiliary constraints.

3 CLOSER: Evaluator-Access Closed-Loop Control

3.1 Problem Setup and Evaluator Roles

We formalize this control cycle for a frozen deterministic sampler with N discrete transitions (Song et al., 2021),

$$z^{(0)} \sim \mathcal{N}(0, I), \qquad z^{(i+1)} = F_i(z^{(i)}; y), \quad i = 0, \dots, N-1, \tag{1}$$

where y is the text condition and the final image is $x = D(z^{(N)})$. CLOSER changes only transient sampler states; the generator, decoder, text encoder, and evaluators remain frozen.

To describe an intervention at position p, write $h = z^{(p)}$ and define the remaining sampler

$$G_{p:N} = F_{N-1} \circ F_{N-2} \circ \cdots \circ F_p. \tag{2}$$

Executing this suffix and decoding its final state gives

$$x_p(h) = D(G_{p:N}(h; y)). \tag{3}$$

The case $p = 0$ recovers initial-noise optimization, while $p > 0$ intervenes on an intermediate state.

The completed image is where feedback is evaluated. We distinguish four roles:

$$J(x, y), \qquad V(x, y), \qquad H(x, y) \in \{0, 1\}, \qquad M(x, y). \tag{4}$$

Here J, V, and M are scores, while H is a binary feasibility test. The differentiable objective J supplies proposal directions; V compares completed candidates, and H applies preservation constraints. These three roles control the trajectory, whereas M reports its final performance. They may use different evaluators, but need not: an evaluator-access protocol may explicitly set $M = V$.

3.2 Exact-Suffix Proposal Gradient

The proposal step connects terminal feedback to the current state. For state h, the differentiable objective is

$$\mathcal{J}_p(h; y) = J(D(G_{p:N}(h; y)), y)\,. \tag{5}$$

Differentiating it with respect to the intervention state gives

$$g_p = \nabla_h \mathcal{J}_p(h; y). \tag{6}$$

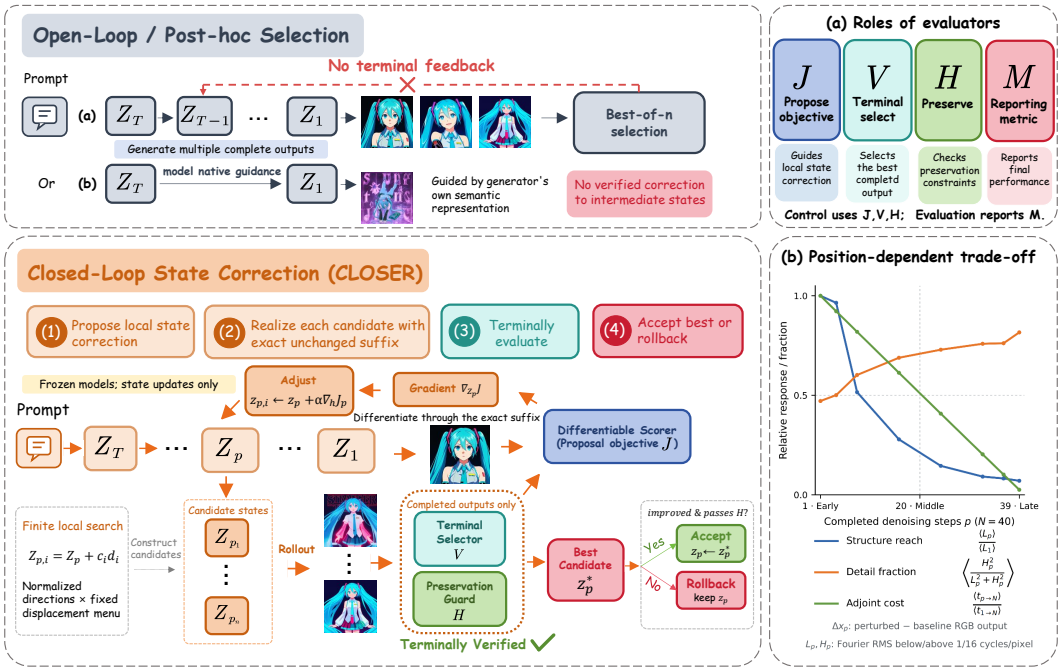


Figure 2: **CLOSER turns a terminal evaluator into a trajectory controller.** The proposal objective J supplies gradients through the remaining sampler. Candidate rollouts then expose the terminal consequences, which selector V and guard H use to accept a correction or retain the incoming state. Reporting metric M has a separate role. Position determines the remaining suffix and thereby influences the scale of the correction.

To keep the proposal tied to the executed trajectory, the differentiation in Eq. (6) uses the sampler, conditioning, CFG setting, time grid, parameterization, solver, and decoder of the candidate rollout. Thus, the derivative follows the implemented discrete suffix. It can be computed by reverse-mode differentiation, starting from

$$\lambda_N = \nabla_{z^{(N)}} J(D(z^{(N)}), y), \tag{7}$$

and propagating backward through its transitions,

$$\lambda_i = \left(DF_i(z^{(i)}; y)\right)^\top \lambda_{i+1}, \qquad i = N-1, \ldots, p, \tag{8}$$

which yields $g_p = \lambda_p$. To reduce memory use, trajectory states can be cached and individual transitions reconstructed during this reverse pass.

3.3 Terminally Verified State Search

The gradient identifies a search direction, but acceptance depends on the completed output. CLOSER therefore first normalizes a nonzero g_p as

$$u_p = \frac{g_p}{\|g_p\|_2 + \epsilon} \tag{9}$$

and uses the normalized direction to construct a finite candidate set

$$\mathcal{C}_p(h) = \{\mathcal{T}_p(h, cd) : d \in \mathcal{D}_p,\ c \in \mathcal{A}_p\}, \tag{10}$$

where $\mathcal{D}_p$ contains one or more normalized evaluator-derived directions, $\mathcal{A}_p$ is a fixed displacement menu, and $\mathcal{T}_p$ is the configured state transport. Intermediate states use $\mathcal{T}_p(h, cd) = h + cd$; an optional seed-level specialization preserves the realized Gaussian radius through spherical transport.

This fixed menu permits instance-dependent choices of direction and magnitude. To assess each choice, CLOSER completes the unchanged suffix,

$$x_p(h') = D(G_{p:N}(h'; y)), \tag{11}$$

and compares the resulting images rather than the state displacements. For margin $\delta \geq 0$, the candidates that satisfy the guard and improve the selector form

$$\mathcal{C}_p^+(h) = \{h' \in \mathcal{C}_p(h) : H(x_p(h'), y) = 1,\ V(x_p(h'), y) \geq V(x_p(h), y) + \delta\}. \tag{12}$$

The best-feasible policy then returns

$$h^\star = \begin{cases} \arg\max\limits_{h' \in \mathcal{C}_p^+(h)} V(x_p(h'), y), & \mathcal{C}_p^+(h) \neq \varnothing, \\ h, & \mathcal{C}_p^+(h) = \varnothing. \end{cases} \tag{13}$$

The second branch retains the incoming state when the search finds no admissible improvement, ensuring selector-relative non-decrease.

Risk-controlled selection. The guard extends this decision to auxiliary criteria. With the incoming image $x = x_p(h)$ as the reference, frozen preservation metrics P_j define

$$H_{\boldsymbol{\tau}}(x', x) = \prod_j \mathbf{1}[P_j(x', y) \geq P_j(x, y) - \tau_j]. \tag{14}$$

These bounds apply relative to the current incoming image. Among feasible candidates within a fixed tolerance of the best target score, selection favors the smallest displacement; if none is feasible, the incoming image is retained.

3.4 Position-Dependent Suffix Reachability

Having defined how to accept a correction, we next consider where to apply it. Position changes the remaining sampler and hence the response to a local intervention δh, which we write as

$$\mathcal{R}_p(h, \delta h) = x_p(h + \delta h) - x_p(h), \qquad \mathcal{R}_p(h, \delta h) \approx D_h x_p(h)\, \delta h. \tag{15}$$

Algorithm 1 Position-Aware CLOSER

Require: Prompt y; frozen sampler $\{F_i\}_{i=0}^{N-1}$; decoder D; stage configurations $\{(p_r, J_r, V_r, H_r, \mathcal{D}_r, \mathcal{A}_r, K_r)\}_{r=1}^{R}$.

```
Sample z^(0) ~ N(0, I) and obtain the baseline trajectory.
for r = 1, ..., R do
    Advance the accepted trajectory to p_r and set h ← z^(p_r).
    for k = 1, ..., K_r do
        Compute the exact-suffix proposal gradient from J_r.
        Construct C_{p_r}(h) from D_r and A_r.
        Complete and evaluate every candidate with V_r and H_r.
        Select the best feasible improvement using Eq. (13).
        if no candidate is accepted then
            break and retain the incoming state.
        else
            Update h to the accepted candidate.
        end if
    end for
end for
Complete the unchanged sampler and return the decoded image.
```

Because the Jacobian depends on the suffix, the same local intervention can have different terminal consequences at different positions. This motivates the frequency-space and perceptual diagnostics in Section 4.4. In the SD3.5M study, early interventions produce broader structural responses, while late interventions favor local refinement, as illustrated in Figure 2.

CLOSER can combine corrections at ordered positions $p_1 \leq \cdots \leq p_R$. Each accepted correction updates the trajectory on which the next gradient is computed. Between these control positions, the original frozen sampler continues without modification.

Algorithm 1 uses best-feasible selection; the OneIG deployment instead scans the fixed menu until the first improvement (Table 9). Positions, evaluator roles, menus, and iteration caps are fixed before evaluation. Acceptance then adapts the direction, displacement, and number of successful corrections to each sample.

4 Experiments

4.1 Experimental Setup

Protocol and benchmarks. We evaluate frozen SD3.5-Medium (SD3.5M) (Esser et al., 2024) and Z-Image (Z-Image Team et al., 2025) using matched prompts, seeds, and deterministic 40-step sampling

at 1024×1024. DPG-Bench (Hu et al., 2024), OneIG-EN (Chang et al., 2025), and GenEval (Ghosh et al., 2023) follow their official protocols; GenEval uses 553 prompts and four seeds. Base/strong CFG are 7/9 for SD3.5M and 4/7 for Z-Image.

Comparison and cost. Within each benchmark, evaluator-access methods use its corresponding terminal evaluator. We compare cost as end-to-end wall-clock normalized by base-CFG generation with the same evaluator ($1.0\times$). Appendix A gives configurations and statistical settings.

4.2 Benchmark Utility and Compute Scaling

Table 1: **In-protocol terminal utility on SD3.5M and Z-Image.** DPG: Overall (%); OneIG: Mean-4; GenEval: four-seed score. **Bold**/underlined: best/second best per backbone.

	SD3.5M			Z-Image		
Method	**DPG ↑**	**OneIG ↑**	**GenEval ↑**	**DPG ↑**	**OneIG ↑**	**GenEval ↑**
Evaluator-free references						
Standard (base CFG) (Ho & Salimans, 2022)	84.99	0.482	0.710	87.47	0.647	0.745
Standard (strong CFG) (Ho & Salimans, 2022)	84.52	0.494	0.702	87.46	0.646	0.765
W2S (Bai et al., 2026)	85.01	0.509	0.699	87.40	0.643	0.782
ZigZag (Bai et al., 2025)	85.67	0.513	0.727	87.29	0.644	0.750
Evaluator-access search and state optimization						
Best-of-4 (Ma et al., 2025)	84.95	0.499	0.745	87.56	0.656	0.782
ZO-3 (Ma et al., 2025)	84.72	0.490	0.742	87.38	0.650	0.766
PSP-4 (Guimaraes & Perona, 2026)	84.98	0.498	0.758	87.29	0.657	0.799
InitNO (Guo et al., 2024)	84.41	0.477	0.679	86.66	0.648	0.665
DOODL (Wallace et al., 2023)	85.04	0.477	0.719	87.64	0.647	0.756
ReNO (Eyring et al., 2024)	84.95	0.476	0.719	87.56	0.649	0.756
CLOSER	**93.05**	**0.526**	**0.871**	**94.30**	**0.663**	**0.905**

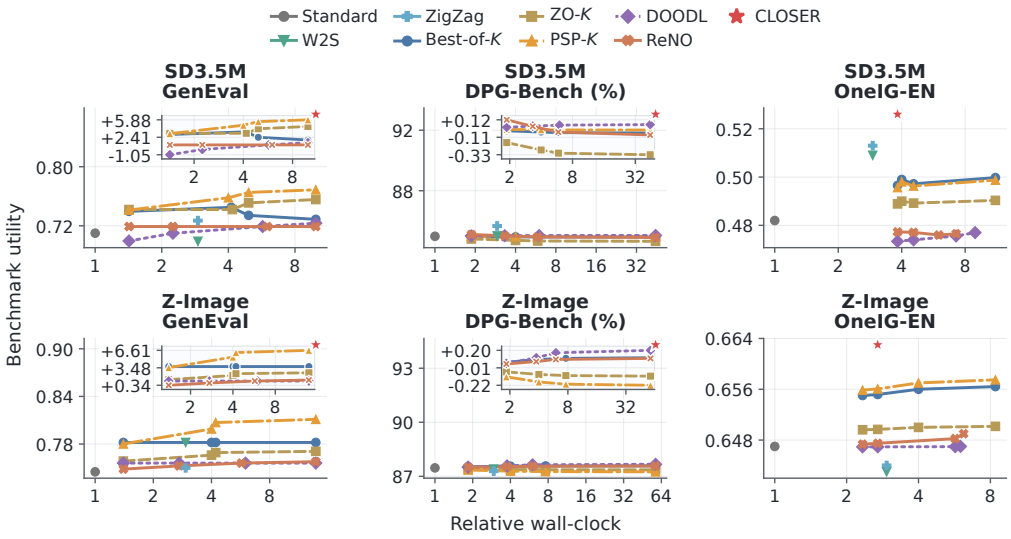


Figure 3: **Compute-matched scaling.** Benchmark utility versus normalized end-to-end wall-clock. The reported CLOSER operating points lie above the swept baseline frontiers in all six settings.

CLOSER ranks first in Table 1, exceeding the strongest baselines by $+7.38/+6.66$ DPG-Bench points and $+0.113/+0.106$ GenEval on SD3.5M/Z-Image. Figure 3 relates these gains to cost through the baseline sweeps; Appendix C gives category breakdowns.

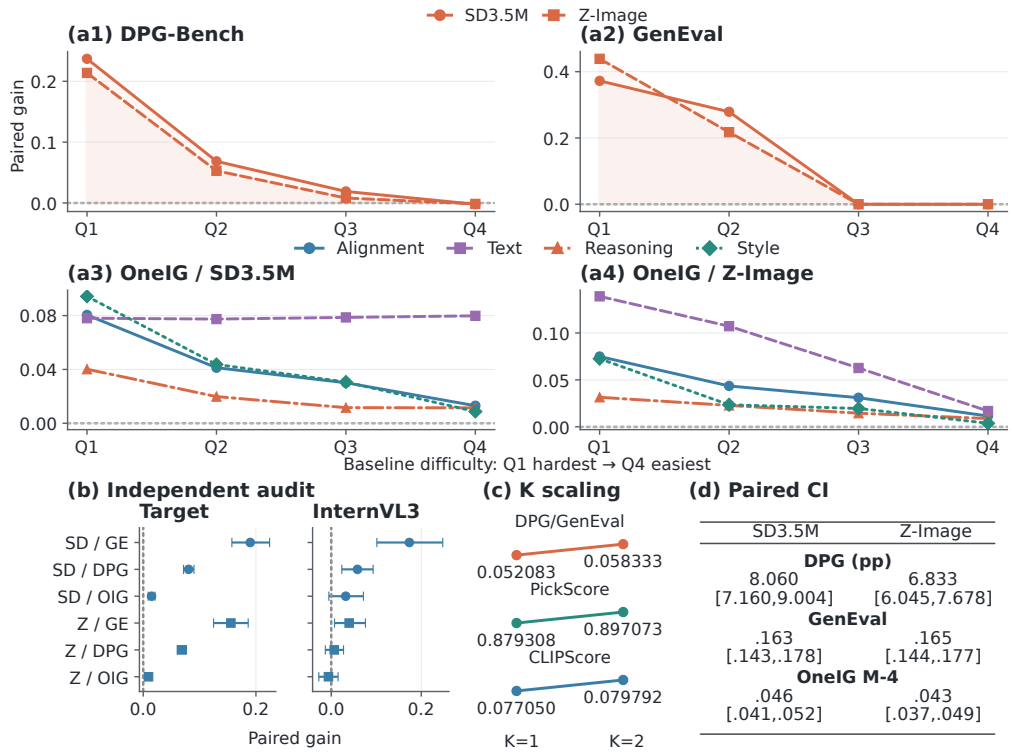


Figure 4: **Distribution, transfer, and finite search. (a1–a4)** Hardest-quartile gains reach $+0.237/+0.214$ on DPG-Bench and $+0.372/+0.439$ on GenEval. **(b)** An unseen InternVL3 judge records three significant transfers and no significant degradation. **(c)** $K = 2$ improves all three selectors in the two-candidate pool. **(d)** All six paired 95% intervals versus Standard lie above zero.

Figure 4**(a1–a4)** locates these gains by baseline difficulty: correction helps most where the incoming output leaves room for improvement. Panel **(b)** tests whether the gains carry beyond the optimized score. To assess their dependence on search and sampling variation, **(c)** adds a second candidate and **(d)** reports paired confidence intervals versus Standard.

4.3 Closed-Loop Mechanism

Table 2: **Closed-loop component ablation at position 1.** The left block reports mean completed-output gain; the right block reports the fraction of prompts whose realized endpoint does not decrease.

Mean completed-output gain				*Non-decreasing endpoint rate*			
Proposal configuration	**Gen.**	**Pick**	**CLIP**	**Acceptance configuration**	**Gen.**	**Pick**	**CLIP**
Random direction	.0708	.1083	.0075	No terminal re-eval.	90.0%	42.5%	47.5%
Single radius	.1958	.2923	.0156	No rollback	97.5%	72.5%	67.5%
Full CLOSER	**.2146**	**.4745**	**.0304**	**Full CLOSER**	**100%**	**100%**	**100%**

Table 2 examines how the gains arise: evaluator-aligned directions and multiple radii improve candidate utility, while re-evaluation and rollback determine which outcomes are retained. Figure 5 extends this distinction to direction controls **(a)**, verification across positions **(b)**, repeated corrections **(c)**, and preservation constraints **(d–f)**.

These controls distinguish task information from step-size choice: random directions test the former, while a single radius removes the latter. The proposal–selector cross-test then holds candidates fixed and changes V for the same J (Table 20). The changed output scores show that terminal choice has an effect beyond candidate construction.

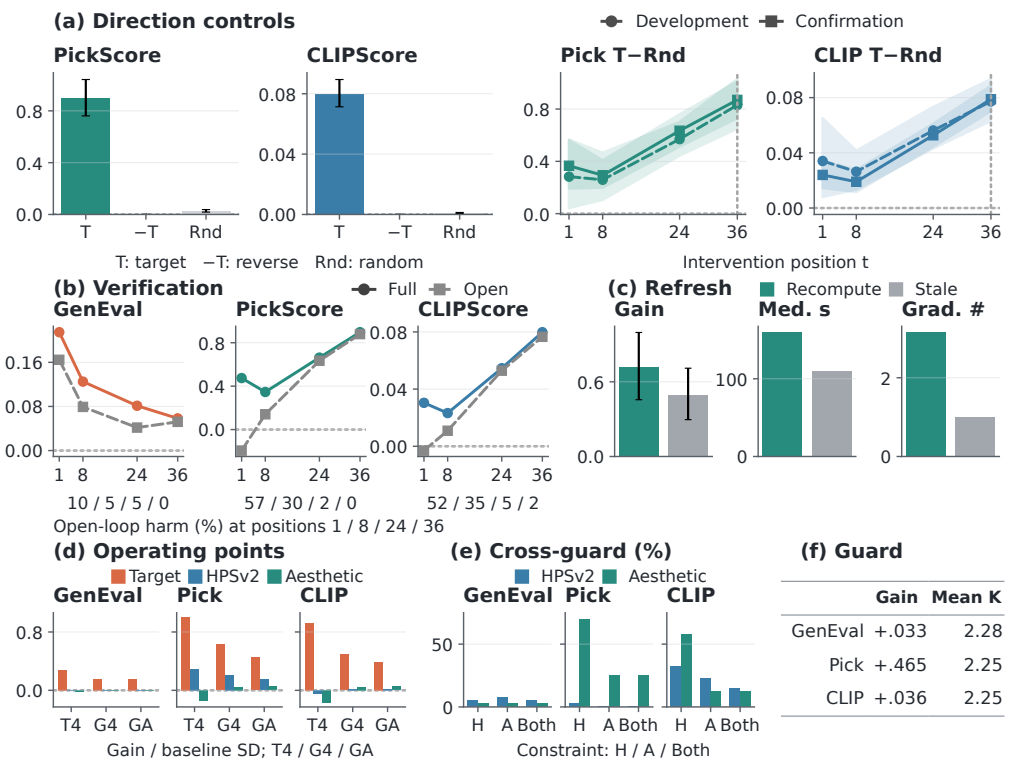


Figure 5: **Closed-loop proposal, verification, and risk control.** **(a)** Target directions outperform reverse and matched-random controls. **(b)** Terminal verification with rollback removes harmful accepted updates at every tested position. **(c)** Proposal refresh raises PickScore gain from $+0.496$ to $+0.722$. **(d–f)** Joint guards preserve both complementary axes while retaining positive target gain with at most 2.28 completed candidates on average.

For repeated corrections, refreshing the gradient makes the next proposal respond to the accepted state rather than the initial trajectory. The guard experiments address a different decision: they hold completed candidates fixed and vary which are admissible. Together with the direction and radius ablations, these tests distinguish candidate quality, sequential adaptation, and preservation at acceptance.

4.4 Intervention Stage Selects the Correction Scale

We next ask where the verified corrections are most effective. On SD3.5M, early intervention favors structure, while late intervention favors preference and alignment (Table 3). Figure 6**(a–c)** relates these utility changes to the frequency and semantic responses in **(d–f)**. Preference/alignment gains rise despite smaller low-frequency changes, indicating that useful correction depends on the response scale, not its magnitude alone. To examine the suffix without terminal selection, random probes provide a separate measurement: normalized low-frequency reach falls from 1.000 to 0.055 and high-frequency share rises from 0.461 to 0.824 (Table 22).

Table 3: **Scorer-selected endpoint stage contrast.** Control shifts from structure to preference/alignment refinement on SD3.5M.

Stage diagnostic	$p = 1$	$p = 36$
Dominant control scale	global	local
Preferred objective	structure	preference/ alignment
Structural-gain advantage	3.7×	–
Preference/alignment gain advantage	–	1.9 × /2.6×

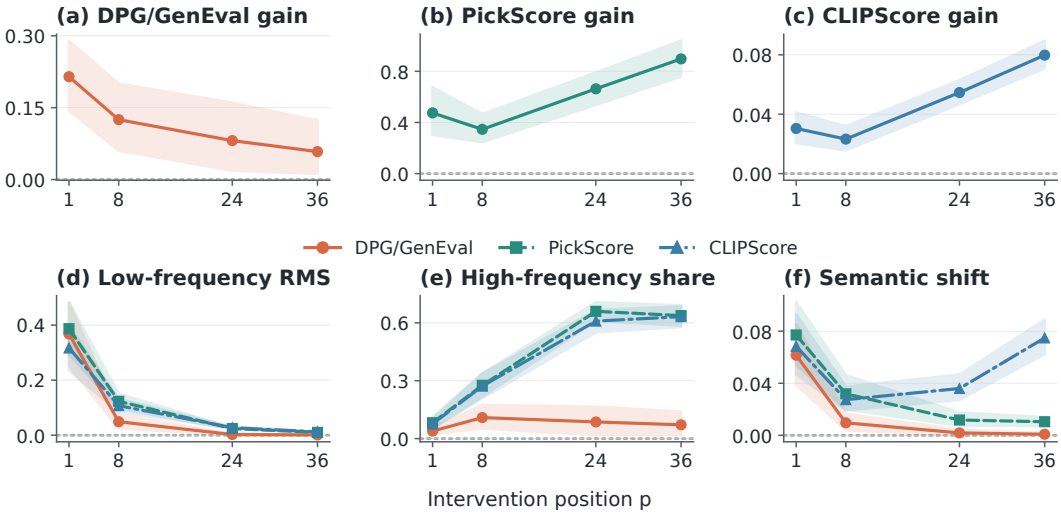


Figure 6: **Objective-dependent stage specialization. (a–c)** In the SD3.5M sweep, structural utility is highest early, while preference and alignment are highest late. **(d–f)** The corresponding selected-output responses shift toward localized, high-frequency refinement as the suffix shortens.

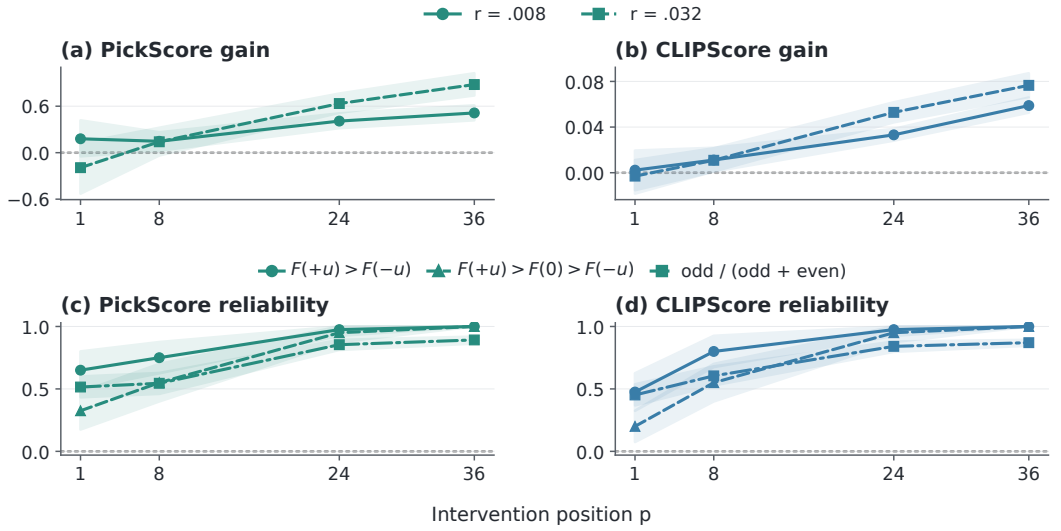


Figure 7: **Locality behind stage specialization.** Target-direction efficacy and local directional reliability rise toward the end of the trajectory, where the larger radius becomes most effective.

Figure 7**(c–d)** distinguishes beating the reverse direction from improving the baseline. The late-stage reliability and shorter suffix costs (Table 19) support local refinement late in the SD3.5M trajectory; deployment positions remain evaluator- and backbone-specific (Table 8).

5 Conclusion

CLOSER turns inference computation into verified corrections of a frozen diffusion trajectory. Gradients propose changes, completed rollouts determine acceptance, and accepted states guide refreshed feedback. This cycle achieves the strongest reported utility in all six settings, with operating points above the compared utility–cost baselines. The SD3.5M diagnostics link correction scale to position: early interventions favor structure and late interventions favor preference and alignment, while random-probe low-frequency reach contracts $10.43\times$ between early and late groups. Together, these results connect inference-time scaling to stage-aware task control without parameter updates.

AI USE STATEMENT

In preparing this work, we used generative AI tools to assist with writing and editing portions of the manuscript and with preliminary research, including literature discovery, summarization, and organization of relevant prior work. All AI-assisted material was reviewed and revised by the authors; citations and factual claims arising from AI-assisted research were checked against the original sources. We take full responsibility for the final content of this submission, including text and research artifacts produced with the assistance of generative AI.

REFERENCES

Lichen Bai, Shitong Shao, Zikai Zhou, Zipeng Qi, Zhiqiang Xu, Haoyi Xiong, and Zeke Xie. Zigzag diffusion sampling: Diffusion models can self-improve via self-reflection. In *International Conference on Learning Representations*, 2025. URL https://iclr.cc/virtual/2025/poster/29938.

Lichen Bai, Masashi Sugiyama, and Zeke Xie. Weak-to-strong diffusion with reflection. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/607864303db1f2bea942b4a9d462920c-Abstract-Conference.html.

Arpit Bansal, Hong-Min Chu, Avi Schwarzschild, Soumyadip Sengupta, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Universal guidance for diffusion models. *arXiv preprint arXiv:2302.07121*, 2023. URL https://arxiv.org/abs/2302.07121.

Jingjing Chang, Yixiao Fang, Peng Xing, Shuhan Wu, Wei Cheng, Rui Wang, Xianfang Zeng, Gang Yu, and Hai-Bao Chen. OneIG-Bench: Omni-dimensional nuanced evaluation for image generation. In *Advances in Neural Information Processing Systems*, volume 38, 2025. URL https://papers.nips.cc/paper_files/paper/2025/hash/e9e9e5428189a3e49479547ef917e88d-Abstract-Datasets_and_Benchmarks_Track.html.

Patrick Esser, Sumith Kulal, Andreas Blattmann, Rahim Entezari, Jonas Müller, Harry Saini, Yam Levi, Dominik Lorenz, Axel Sauer, Frederic Boesel, Dustin Podell, Tim Dockhorn, Zion English, and Robin Rombach. Scaling rectified flow transformers for high-resolution image synthesis. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12606–12633. PMLR, 2024. URL https://proceedings.mlr.press/v235/esser24a.html.

Luca Eyring, Shyamgopal Karthik, Karsten Roth, Alexey Dosovitskiy, and Zeynep Akata. ReNO: Enhancing one-step text-to-image models through reward-based noise optimization. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-3987. URL https://papers.nips.cc/paper_files/paper/2024/hash/e31bdea0a93741c2157eea705dd219eb-Abstract-Conference.html.

Dhruba Ghosh, Hannaneh Hajishirzi, and Ludwig Schmidt. GenEval: An object-focused framework for evaluating text-to-image alignment. *arXiv preprint arXiv:2310.11513*, 2023. URL https://arxiv.org/abs/2310.11513.

Rogerio Guimaraes and Pietro Perona. Inference-time scaling of diffusion models via progressive seed pruning. *arXiv preprint arXiv:2607.21591*, 2026. URL https://arxiv.org/abs/2607.21591.

Xiefan Guo, Jinlin Liu, Miaomiao Cui, Jiankai Li, Hongyu Yang, and Di Huang. InitNO: Boosting text-to-image diffusion models via initial noise optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9380–9389, 2024. URL https://openaccess.thecvf.com/content/CVPR2024/html/Guo_InitNO_Boosting_Text-to-Image_Diffusion_Models_via_Initial_Noise_Optimization_CVPR_2024_paper.html.

Jack Hessel, Ari Holtzman, Maxwell Forbes, Ronan Le Bras, and Yejin Choi. CLIPScore: A reference-free evaluation metric for image captioning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 7514–7528. Association

for Computational Linguistics, 2021. doi: 10.18653/v1/2021.emnlp-main.595. URL https://aclanthology.org/2021.emnlp-main.595/.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022. URL https://arxiv.org/abs/2207.12598.

Xiwei Hu, Rui Wang, Yixiao Fang, Bin Fu, Pei Cheng, and Gang Yu. ELLA: Equip diffusion models with LLM for enhanced semantic alignment. *arXiv preprint arXiv:2403.05135*, 2024. URL https://arxiv.org/abs/2403.05135.

Yuval Kirstain, Adam Polyak, Uriel Singer, Shahbuland Matiana, Joe Penna, and Omer Levy. Pick-a-pic: An open dataset of user preferences for text-to-image generation. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-1594. URL https://papers.nips.cc/paper_files/paper/2023/hash/73aacd8b3b05b4b503d58310b523553c-Abstract-Conference.html.

Nanye Ma, Shangyuan Tong, Haolin Jia, Hexiang Hu, Yu-Chuan Su, Mingda Zhang, Xuan Yang, Yandong Li, Tommi Jaakkola, Xuhui Jia, and Saining Xie. Inference-time scaling for diffusion models beyond scaling denoising steps. *arXiv preprint arXiv:2501.09732*, 2025. URL https://arxiv.org/abs/2501.09732.

Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=St1giarCHLP.

Bram Wallace, Akash Gokul, Stefano Ermon, and Nikhil Naik. End-to-end diffusion latent optimization improves classifier guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 7280–7290, 2023. URL https://openaccess.thecvf.com/content/ICCV2023/html/Wallace_End-to-End_Diffusion_Latent_Optimization_Improves_Classifier_Guidance_ICCV_2023_paper.html.

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

Jiazheng Xu, Xiao Liu, Yuchen Wu, Yuxuan Tong, Qinkai Li, Ming Ding, Jie Tang, and Yuxiao Dong. ImageReward: Learning and evaluating human preferences for text-to-image generation. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-0700. URL https://papers.nips.cc/paper_files/paper/2023/hash/33646ef0ed554145eab65f6250fab0c9-Abstract-Conference.html.

Haotian Ye, Haowei Lin, Jiaqi Han, Minkai Xu, Sheng Liu, Yitao Liang, Jianzhu Ma, James Zou, and Stefano Ermon. TFG: Unified training-free guidance for diffusion models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-0704. URL https://papers.nips.cc/paper_files/paper/2024/hash/2818054fc6de6dacdda0f142a3475933-Abstract-Conference.html.

Jiwen Yu, Yinhuai Wang, Chen Zhao, Bernard Ghanem, and Jian Zhang. FreeDoM: Training-free energy-guided conditional diffusion model. *arXiv preprint arXiv:2303.09833*, 2023. URL https://arxiv.org/abs/2303.09833.

Z-Image Team, Huanqia Cai, Sihan Cao, Ruoyi Du, Peng Gao, Steven Hoi, Shijie Huang, Zhaohui Hou, Dengyang Jiang, Xin Jin, Liangchen Li, Zhen Li, Zhong-Yu Li, David Liu, Dongyang Liu, Junhan Shi, Qilong Wu, Feng Yu, Chi Zhang, Shifeng Zhang, and Shilin Zhou. Z-Image: An efficient image generation foundation model with single-stream diffusion transformer. *arXiv preprint arXiv:2511.22699*, 2025. URL https://arxiv.org/abs/2511.22699.

A Reproducibility Settings

We first specify the shared generation protocol, then the benchmark-specific CLOSER configurations, evaluator routing, baseline settings, and statistical procedures. Candidate menus, intervention positions, and iteration caps are fixed before test evaluation. Within these fixed configurations, the accepted displacement and number of successful corrections depend on the instance.

A.1 Shared Generation and Benchmark Protocols

Table 4: **Shared generation settings.**

Setting	SD3.5M	Z-Image
Resolution	1024×1024	1024×1024
Sampler	deterministic FlowMatch Euler	deterministic FlowMatch Euler
Denoising steps	40	40
CFG / guidance	7	4
Negative prompt	empty	empty
Maximum text length	512	512
Denoiser / VAE compute	BF16	BF16
Persistent sampler state	FP32	FP32
Conditioning	pre-encoded, fixed	pre-encoded, fixed
Stochastic suffix noise	none	none

Under these shared settings, paired Standard–CLOSER comparisons use the same prompt and initial seed. Resetting the scheduler before every baseline or candidate trajectory keeps the continuation settings matched.

Table 5: **Benchmark sample counts and reporting protocol.**

Benchmark	Examples	Seeds	Reported score
GenEval	553	42–45	official task scores and macro overall
DPG-Bench	1,065	42	official hard-DPG categories and overall
OneIG-EN	1,120	42	Alignment, Text, Reasoning, Style, Mean-4

For benchmark-scale GenEval, each of the 553 prompts is evaluated with seeds 42–45, and resampling keeps its four outputs in one cluster. DPG-Bench and OneIG-EN instead use one prompt/example as the statistical unit. The reported OneIG Mean-4 is the average of its four component scores.

A.2 Frozen CLOSER Configurations

Table 6: **GenEval CLOSER hyperparameters.**

Hyperparameter	SD3.5M	Z-Image
Intervention position (0-based)	1	4
Maximum correction iterations	4	2
Displacement menu	$\{2\}$	$\{2, 1, 0.5, 0.25\}$
Proposal surrogate directions	Pick, CLIP, normalized Pick+CLIP consensus	same
Direction signs	positive and negative	same
Terminal objective	$\Delta S_{\text{strict}} + 0.25\Delta S_{\text{dense}}$	same
Guard	GenEval non-decrease	same
Improvement threshold	0	0
Candidate selection	best feasible improving candidate	same
Early stop	strict GenEval success	same
Rollback	enabled	enabled

Table 7: **DPG-Bench CLOSER hyperparameters.**

Hyperparameter	SD3.5M	Z-Image
Early intervention position	1	0
Early maximum iterations	8	8
Early proposal objective	terminal soft DPG	terminal soft DPG
Early displacement menu	$\{32, 16, 8, 4, 2, 1, 0.5, 0.25\}$	same
Early selection	best improving	best improving
Minimum improvement	10^{-5}	10^{-5}
Gradient recomputation	after every accepted iteration	same
Middle intervention position	32	32
Middle maximum iterations	4	4
Middle proposal objective	terminal soft DPG	terminal soft DPG
Middle displacement menu	$\{2, 1, 0.5, 0.25\}$	$\{8, 4, 2, 1, 0.5, 0.25\}$
Middle selection	best improving	best improving
DPG preservation floor	$\max(S_{\text{base}}, S_{\text{early}} - 0.05)$	same
DPG dependency mode	soft	soft
Low-score log weight	0	0
DPG question batch	8	8
Soft-DPG compute	BF16	BF16
Rollback	enabled	enabled

OneIG-EN search and deployment. OneIG uses the same intermediate-state CLOSER mechanism as the other main benchmarks. The frozen position search uses 10 examples per backbone, candidate positions $\{0, 8, 24, 36\}$, one correction iteration, the ordered displacement menu $\{16, 8, 4, 2, 1, 0.5, 0.25\}$, and first-improving selection with threshold 0. Route-wise mean terminal-objective gain selects the frozen position, with acceptance rate used to break ties.

Table 8: **Frozen OneIG intervention positions.**

Evaluator route	SD3.5M	Z-Image
Alignment	36	0
Alignment + Style	36	0
Text	0	24
Reasoning	0	8

Table 9: **Shared OneIG deployment hyperparameters.**

Hyperparameter	Value
Controlled variable	sampler state at the frozen route-specific position
Maximum iterations	1
Ordered displacement menu	$\{16, 8, 4, 2, 1, 0.5, 0.25\}$
Proposal direction	normalized exact terminal OneIG suffix gradient
Direction signs	positive only
Selection policy	first improving
Improvement threshold	0
Candidate evaluation	exact remaining suffix, decoder, and route evaluator
Rollback	baseline when no candidate improves
Component weights	1 for every active component
Component-gradient normalization	disabled
Gradient checkpointing	enabled
Qwen image input size	224
Qwen vision branch precision	FP32

With these deployment settings, the prompt category determines the OneIG route: `object`, `human`, and `multilingualism` use Alignment; `anime` uses Alignment+Style when a style class is available; `text` uses Text; and `reasoning` uses Reasoning. This assigns the 1,120 examples to 581 Alignment-only, 114 Alignment+Style, 200 Text, and 225 Reasoning routes.

A.3 Evaluator Mapping

Table 10: **Benchmark-matched terminal evaluators.** Proposal surrogates construct candidates; terminal selection and reporting use the evaluator corresponding to each benchmark.

Benchmark / route	Proposal signal	Terminal evaluator	Final report
GenEval (Ghosh et al., 2023)	PickScore (Kirstain et al., 2023) / CLIPScore (Hessel et al., 2021) proposal surrogates	official GenEval	official GenEval
DPG-Bench (Hu et al., 2024)	differentiable soft-DPG surrogate	DPG evaluator with preservation floor	official hard DPG
OneIG Alignment	differentiable Qwen alignment	OneIG Alignment	OneIG Alignment
OneIG Text	differentiable Qwen text likelihood	OneIG Text	OneIG Text
OneIG Reasoning	differentiable LLM2CLIP	OneIG Reasoning	OneIG Reasoning
OneIG Style	differentiable CSD/style	OneIG Style	OneIG Style

For the proposal computations in Table 10, image paths remain floating point from VAE output through resizing and normalization.

A.4 Baseline Hyperparameters

Table 11: **Frozen settings for reproduced baselines.**

Baseline	Frozen setting
CFG Standard (Ho & Salimans, 2022)	SD3.5M CFG 7; Z-Image CFG 4.
CFG Large (Ho & Salimans, 2022)	SD3.5M CFG 9; Z-Image CFG 7.
W2S (Bai et al., 2026)	strong/weak CFG $(9, 1)$ for SD3.5M and $(7, 1)$ for Z-Image.
ZigZag (Bai et al., 2025)	forward–inverse–forward correction over the first 39 steps; inverse CFG 0.
Best-of-4 (Ma et al., 2025)	four independent complete trajectories; terminal reranking with the evaluator corresponding to the benchmark.
ZO-3 (Ma et al., 2025)	pivot plus three radius-preserving neighbors; cosine similarity 0.95; one round; terminal acceptance with the evaluator corresponding to the benchmark.
PSP-4 (Guimaraes & Perona, 2026)	eight candidates to 25% progress, retain four; to 50%, retain two; finish and rerank with the evaluator corresponding to the benchmark.
InitNO (Guo et al., 2024)	conditional–unconditional velocity probe at $t^\star = 900$; displacement 12; radius retraction.
DOODL (Wallace et al., 2023)	one 40-step optimization update using the benchmark-matched differentiable objective; spherical coefficient 0.1; 16 crops; cut power 0.3; momentum 0.9; element clip 10^{-3}; noise std. 10^{-4}.
ReNO (Eyring et al., 2024)	one 40-step Nesterov update using the benchmark-matched differentiable objective; learning rate 5; momentum 0.9; gradient clip 0.1; regularizer 0.01.

The evaluator-access baselines share benchmark routing: official GenEval, DPG, or the corresponding OneIG component. Their trajectory counts and update rules differ as recorded in Table 11. To compare these different operations on a common cost axis, reported inference costs use the normalized end-to-end wall-clock convention in Section 4.1.

A.5 Statistical Settings

Because candidates from the same prompt share the search context, statistical resampling is performed at the prompt level rather than over candidate images.

Table 12: **Statistical settings for paired analyses.**

Setting	Value
Bootstrap replicates	10,000
Confidence level	95%
Bootstrap seed	20260823
Resampling unit	prompt cluster
Pairing	identical prompt and initial seed
Win/tie threshold	10^{-12} absolute delta
Tail statistic	worst 10% mean

B Implementation and Evaluation Details

B.1 Deterministic Sampler Replay

A common deterministic continuation makes candidate outcomes comparable. For each prompt and initial seed, we execute the frozen 40-step sampler and cache the states needed at the configured control positions. A candidate at p keeps the executed prefix, replaces $z^{(p)}$, and runs the remaining suffix with the same condition, CFG scale, timestep grid, solver, and decoder. Differences between candidates therefore reflect state interventions rather than resampled solver noise.

For proposal construction, only the selected state is marked as differentiable: model parameters remain frozen and the suffix gradient is evaluated in FP32. Candidate evaluation then disables gradient tracking. When multiple corrections are enabled, each accepted update triggers a new suffix-gradient computation at the updated state, following Algorithm 1.

B.2 Risk-Controlled Selection Protocol

To isolate the effect of preservation at selection, the risk-controlled experiment keeps a pool of completed suffix rollouts fixed and re-selects from it under different guards. For auxiliary score P_j, the floor is defined relative to the incoming image as $P_j(x') \geq P_j(x) - 0.05\sigma_j$, where σ_j is the baseline standard deviation over the analysis set. The two registered auxiliary axes are HPSv2 and Aesthetic score. If several feasible candidates are within 0.1 target standard deviations of the best feasible target score, the smallest intervention radius is selected. The adaptive menu assigns $K \in \{4, 2, 1\}$ to low-, medium-, and high-baseline-score strata and uses 2.25–2.28 candidates on average. The incoming image remains available as the fallback in every configuration.

Table 13: **Full target-only and risk-controlled selection results.** All policies re-select from the same frozen pool of completed suffix rollouts.

Proposal	Selection	Target	HPSv2	Aesthetic	Accept	Mean K
GenEval	Target-only $K = 4$	+0.0583	-0.0320	-0.0093	12.5%	4.00
	Guarded $K = 4$	+0.0333	-0.0045	-0.0002	5.0%	4.00
	Guarded adaptive	+0.0333	-0.0045	-0.0002	5.0%	2.28
PickScore	Target-only $K = 4$	+1.0442	+1.0269	-0.0509	100%	4.00
	Guarded $K = 4$	+0.6498	+0.7458	+0.0170	77.5%	4.00
	Guarded adaptive	+0.4649	+0.5256	+0.0187	65.0%	2.25
CLIPScore	Target-only $K = 4$	+0.0851	-0.1529	-0.0601	100%	4.00
	Guarded $K = 4$	+0.0455	+0.0570	+0.0173	62.5%	4.00
	Guarded adaptive	+0.0362	+0.0558	+0.0184	52.5%	2.25

Under these constraints, the adaptive guard retains positive target gain for all three selectors with 2.25–2.28 candidates on average. For PickScore and CLIPScore, both registered auxiliary scores also show positive mean changes.

B.3 Intervention Positions and Candidate Search

Within the 40-transition sampler, the position study evaluates $p \in \{1, 8, 24, 36\}$ to compare early, intermediate, and late corrections. Benchmark-specific deployment positions are then fixed before final evaluation. At each tested position, the normalized terminal gradient defines the proposal direction and a finite displacement menu defines the candidates. These are compared with the incoming state, which remains the rollback option.

The closed-loop ablation first compares one displacement with a two-radius menu. The budget experiments extend this idea using nested menus within each saved pool: every larger menu contains the smaller menu's candidates. Keeping prompts, seeds, directions, and solver settings fixed within a pool isolates the contribution of additional completed suffix evaluations.

B.4 Result Families and Estimands

Benchmark-scale performance and controlled mechanism studies answer different questions and accordingly use different prompt sets, candidate pools, configurations, and aggregation rules. Table 14 summarizes the setting, statistical unit, and estimand associated with each result family.

Table 14: **Experimental settings and reported estimands.** The table lists the data split, controller configuration, statistical unit, and reported estimand for each result family.

Result family	Data, generation, and controller settings	Statistical unit and estimand
Main DPG-Bench table	Full 1,065-prompt benchmark; one matched seed per prompt; 1024^2; deterministic 40-step sampler Benchmark-specific final CLOSER configuration	Official benchmark aggregation Official hard-DPG Overall score
Main GenEval table	Official 553 prompts $\times$ four matched seeds; 1024^2; deterministic 40-step sampler Benchmark-specific final CLOSER configuration	Official category/seed aggregation Official GenEval aggregate
Main OneIG-EN table	Full English split with matched initial seeds; 1024^2; deterministic 40-step sampler Component-specific final CLOSER configuration	Official component aggregation Arithmetic mean of the four official component scores
Paired main-result analysis	Matched Standard (base CFG) and CLOSER outputs from the benchmark runs above Same final configurations as the corresponding benchmark run	Prompt; repeated seeds clustered within prompt Mean prompt-paired score change relative to Standard (base CFG), with a prompt-bootstrap interval
Cross-evaluator and risk-control studies	SD3.5M; 40 disjoint confirmation prompts; frozen pool of completed suffix rollouts, primarily at position 36 Common proposal directions/radii; selectors and guards re-select the same pool	Prompt Mean paired terminal-score change for the specified selector/reporting metric
Closed-loop component ablation	SD3.5M; 40 confirmation prompts; position 1; matched completed candidate pools Full loop versus random direction, single radius, no terminal re-evaluation, or no rollback	Prompt Mean active-selector gain and non-decreasing endpoint rate
Position and reachability analysis	SD3.5M; 20 development prompts and 40 disjoint confirmation prompts; positions $\{1, 8, 24, 36\}$ Scorer-selected two-radius outputs; separate unselected signed random probes	Prompt Selected-output gains and response diagnostics; separate reward-independent random reach (Appendix D.6)
Gradient-refresh study	SD3.5M; 20 paired confirmation prompts; up to four correction rounds Fresh gradient after acceptance versus reuse of the initial gradient	Prompt Paired difference in final PickScore gain

In particular, Table 1 follows each benchmark's released aggregation rule, while Figure 4(d) reports the mean within-prompt change from Standard (base CFG) to CLOSER. The latter supports paired bootstrap inference over prompts and is a distinct estimand from the official aggregate difference.

C Detailed Quantitative Results

C.1 DPG-Bench Breakdown

Table 15: **Detailed DPG-Bench results.** All values are percentages. CLOSER is highlighted in bold.

	SD3.5M						Z-Image					
Method	Glo.	Ent.	Attr.	Rel.	Other	Overall	Glo.	Ent.	Attr.	Rel.	Other	Overall
Standard (base CFG) (Ho & Salimans, 2022)	83.89	90.45	88.48	93.69	82.40	84.99	85.41	92.37	90.06	94.39	89.60	87.47
Standard (strong CFG) (Ho & Salimans, 2022)	82.98	90.50	88.18	93.46	81.60	84.52	85.11	92.50	90.44	94.82	88.00	87.46
W2S (Bai et al., 2026)	82.67	90.61	88.34	93.89	83.60	85.01	84.19	92.37	89.94	94.93	88.80	87.40
ZigZag (Bai et al., 2025)	82.67	90.81	88.74	93.85	85.60	85.67	86.32	92.37	89.46	94.78	89.60	87.29
Best-of-4 (Ma et al., 2025)	84.50	90.29	88.40	93.11	86.00	84.95	84.50	92.52	90.48	94.70	90.80	87.56
ZO-3 (Ma et al., 2025)	83.89	90.19	88.62	93.58	82.80	84.72	82.37	92.29	89.92	94.43	89.20	87.38
PSP-4 (Guimaraes & Perona, 2026)	85.71	90.50	88.01	93.50	83.20	84.98	85.11	92.31	90.16	94.74	89.20	87.29
InitNO (Guo et al., 2024)	86.32	89.95	88.34	94.08	80.00	84.41	86.02	91.92	90.02	94.31	86.00	86.66
DOODL (Wallace et al., 2023)	85.11	90.39	88.22	93.58	83.60	85.04	84.50	92.21	90.26	94.70	87.60	87.64
ReNO (Eyring et al., 2024)	85.11	90.24	88.64	93.97	82.40	84.95	85.11	92.29	89.92	94.70	88.40	87.56
CLOSER	**90.58**	**95.48**	**92.81**	**96.21**	**90.80**	**93.05**	**91.19**	**96.59**	**94.56**	**96.40**	**95.20**	**94.30**

Table 16: **Fine-grained DPG-Bench category breakdown on SD3.5M.** Scores use the official hard-DPG level-2 categories under the same 1024×1024 protocol. All values are percentages.

Method	**Attribute**					**Entity**			**Other**		**Relation**	
	Color	Other	Shape	Size	Texture	Part	State	Whole	Count	Text	Non-spatial	Spatial
Standard (CFG=7) (Ho & Salimans, 2022)	92.93	87.36	80.35	70.66	87.48	84.57	83.10	92.56	79.50	94.00	90.57	93.90
CLOSER	**95.68**	**91.13**	**89.52**	**81.82**	**92.32**	**91.41**	**89.65**	**97.08**	**89.00**	**98.00**	**92.45**	**96.46**
Δ	**+2.76**	**+3.77**	**+9.17**	**+11.16**	**+4.84**	**+6.84**	**+6.55**	**+4.52**	**+9.50**	**+4.00**	**+1.89**	**+2.56**

C.2 OneIG-EN Breakdown

Table 17: **Detailed OneIG-EN results.** Mean-4 averages Alignment, Text, Reasoning, and Style.

	SD3.5M					Z-Image				
Method	Align.	Text	Reason.	Style	Mean-4	Align.	Text	Reason.	Style	Mean-4
Standard (base CFG) (Ho & Salimans, 2022)	.801	.493	.288	.347	.482	.868	.981	.299	.439	.647
Standard (strong CFG) (Ho & Salimans, 2022)	.809	.530	.293	.345	.494	.872	.982	.302	.428	.646
W2S (Bai et al., 2026)	.797	.587	.303	.351	.509	.864	.985	.306	.415	.643
ZigZag (Bai et al., 2025)	.800	.591	.304	.356	.513	.867	.988	.306	.414	.644
Best-of-4 (Ma et al., 2025)	.807	.544	.296	.350	.499	.875	.982	.312	.455	.656
ZO-3 (Ma et al., 2025)	.803	.543	.299	.314	.490	.868	.979	.309	.446	.650
PSP-4 (Guimaraes & Perona, 2026)	.809	.542	.302	.340	.498	.874	.987	.316	.453	.657
InitNO (Guo et al., 2024)	.800	.483	.287	.338	.477	.870	.972	.290	.460	.648
DOODL (Wallace et al., 2023)	.802	.473	.288	.344	.477	.868	.979	.295	.446	.647
ReNO (Eyring et al., 2024)	.802	.463	.292	.348	.476	.871	.978	.304	.441	.649
CLOSER	**.821**	**.584**	**.308**	**.391**	**.526**	**.878**	**.988**	**.318**	**.469**	**.663**

C.3 GenEval Breakdown

Table 18: **Detailed GenEval results.** The official aggregate score is reported in the final column of each backbone block. All methods are evaluated with the official 553-prompt, four-seed GenEval protocol.

	SD3.5M							Z-Image						
Method	Single	Two	Count	Color	Pos.	Bind	Overall	Single	Two	Count	Color	Pos.	Bind	Overall
Standard (base CFG) (Ho & Salimans, 2022)	.997	.881	.694	.832	.265	.593	.710	.997	.907	.731	.883	.403	.548	.745
Standard (strong CFG) (Ho & Salimans, 2022)	1.000	.869	.738	.809	.250	.550	.702	1.000	.929	.800	.872	.430	.560	.765
W2S (Bai et al., 2026)	.988	.909	.688	.798	.300	.510	.699	.988	.899	.813	.883	.470	.640	.782
ZigZag (Bai et al., 2025)	.988	.909	.788	.819	.290	.570	.727	1.000	.909	.750	.883	.420	.540	.750
Best-of-4 (Ma et al., 2025)	1.000	.939	.788	.840	.280	.620	.745	.988	.909	.863	.872	.440	.620	.782
ZO-3 (Ma et al., 2025)	.988	.889	.838	.798	.280	.660	.742	1.000	.889	.763	.915	.410	.620	.766
PSP-4 (Guimaraes & Perona, 2026)	1.000	.929	.838	.872	.290	.620	.758	1.000	.889	.900	.872	.460	.670	.799
InitNO (Guo et al., 2024)	.988	.798	.713	.809	.190	.580	.679	.888	.859	.575	.766	.370	.530	.665
DOODL (Wallace et al., 2023)	1.000	.869	.775	.809	.240	.620	.719	.988	.889	.788	.851	.460	.560	.756
ReNO (Eyring et al., 2024)	1.000	.889	.738	.830	.230	.630	.719	1.000	.899	.750	.904	.420	.560	.756
CLOSER	**1.000**	**.982**	**.959**	**.915**	**.503**	**.865**	**.871**	**1.000**	**.977**	**.959**	**.973**	**.645**	**.875**	**.905**

C.4 Qualitative Comparison

The qualitative examples complement the category and difficulty analyses by showing corrections to spatial relations, object counts, fine-grained attributes, and dense scene content. Within each comparison, the generator remains frozen and the initial prompt–seed pairing is unchanged.



Figure 8: **Qualitative comparison on SD3.5M and Z-Image.** Methods within each row share the same prompt and initial seed. The examples cover spatial layout, dense attributes, counting, and long-form scene composition.

D Additional Closed-Loop Analyses

D.1 Position-Wise Interpretation

The position sweeps in Figures 5 and 6 separate the effects of placement and acceptance. Placement changes the remaining suffix and the terminal response it produces; verification decides which realized response is retained. In the SD3.5M study, early proposals pass through longer suffixes and produce broader structural changes, whereas late proposals favor local refinement. Terminal accept-or-rollback selection operates at both ends of this spectrum.

D.2 Candidate Search and Gradient Refresh

Candidate search varies how many completed alternatives are available: Figure 4(c) uses a two-candidate pool at $K = 1, 2$, while Table 21 uses a separately generated eight-candidate pool at $K = 1, 2, 4, 8$. Appendix D.6 defines both menus and their paired statistics. By contrast, the refresh comparison in Figure 5(c) varies whether subsequent proposals use the accepted state or reuse the initial gradient.

D.3 Unified Inference-Cost Profile

Table 19: **CLOSER inference-cost profile.** Inference cost is measured as end-to-end wall-clock latency and normalized by Standard (base CFG) 40-step generation with the benchmark-matched terminal evaluator ($1.0\times$), following the convention used throughout the paper.

Position	Objective	Relative inference cost	Target gain
1	GenEval	$4.52\times$	+0.2146
1	PickScore	$6.68\times$	+0.4745
1	CLIPScore	$6.68\times$	+0.0304
36	GenEval	$1.45\times$	+0.0583
36	PickScore	$1.66\times$	+0.8971
36	CLIPScore	$1.66\times$	+0.0798

The same wall-clock normalization links two levels of cost analysis. Figure 3 compares benchmark-scale operating points, whereas Table 19 examines how the differentiable suffix length changes the cost of a fixed two-radius control configuration.

At position 36, this configuration costs $1.45\times$ for GenEval and $1.66\times$ for PickScore or CLIPScore. Moving to position 1 extends the suffix and raises the measured cost to $4.52\times$–$6.68\times$. Both measurements include sampling, suffix differentiation, candidate rollouts, decoding, and evaluator execution.

D.4 Transfer and Proposal–Selector Decoupling

Table 20 crosses proposal objectives with terminal selectors. For each proposal objective, its completed candidates are held fixed while the selector changes. This separates the influence of the selection criterion from that of candidate construction.

Table 20: **Proposal–selector decoupling at position 36.**

J	V	ΔGenEval	ΔPick	ΔCLIP	ΔHPSv2	ΔAesthetic
Pick	Pick	+0.019	+0.897	+0.004	+0.849	-0.050
Pick	CLIP	+0.019	+0.594	+0.008	+0.560	-0.033
CLIP	Pick	+0.000	+0.051	+0.042	+0.035	-0.015
CLIP	CLIP	+0.025	-0.060	+0.080	-0.085	-0.038

A complementary test holds the selector fixed and increases the candidate budget. Table 21 and Figure 9 report both target gains and changes in HPSv2 and Aesthetic, which remain outside proposal construction and selection in this experiment.

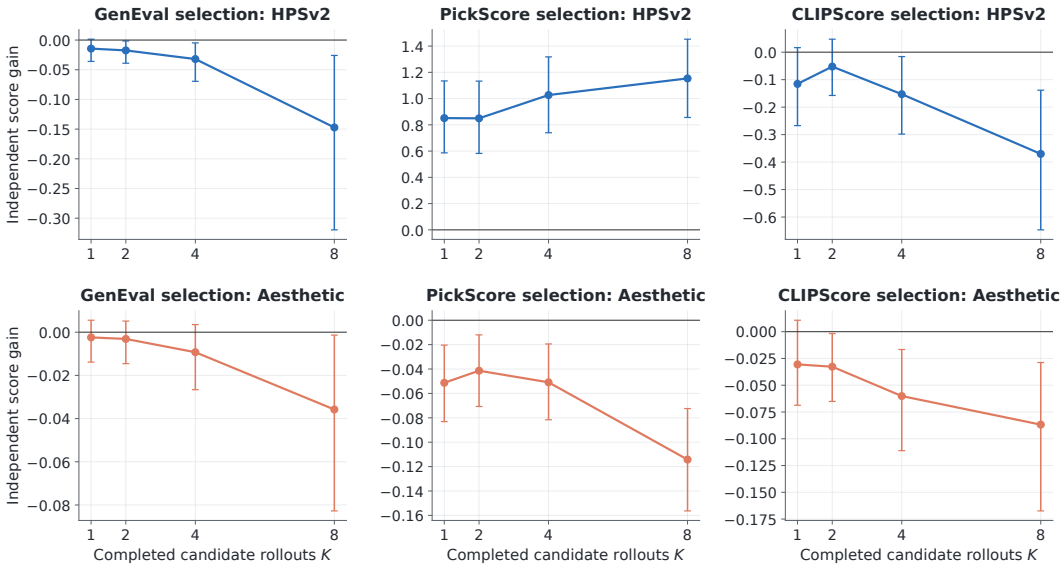


Figure 9: **Independent transfer as completed candidate budget grows.** Error bars are prompt-level bootstrap 95% confidence intervals.

Table 21: **Candidate-budget scaling and independent transfer.** All values are mean paired gains; in this experiment, HPSv2 and Aesthetic are not used for proposal construction or selection.

Selector	K	Target	HPSv2	Aesthetic	Accept
GenEval	1	+0.0458	-0.0144	-0.0024	10.0%
	2	+0.0458	-0.0175	-0.0031	10.0%
	4	+0.0583	-0.0320	-0.0093	12.5%
	8	+0.0958	-0.1472	-0.0358	20.0%
PickScore	1	+0.8925	+0.8513	-0.0513	97.5%
	2	+0.9090	+0.8497	-0.0414	100%
	4	+1.0442	+1.0269	-0.0509	100%
	8	+1.1340	+1.1537	-0.1142	100%
CLIPScore	1	+0.0764	-0.1155	-0.0306	97.5%
	2	+0.0796	-0.0524	-0.0327	100%
	4	+0.0851	-0.1529	-0.0601	100%
	8	+0.0874	-0.3702	-0.0868	100%

D.5 Seed-Level Diversity

To assess variation across seeds, we compute all six pairwise LPIPS distances among the four matched GenEval outputs for each of 553 prompts. The mean changes from 0.5986 to 0.6021 on SD3.5M and from 0.6700 to 0.6692 on Z-Image, indicating stable within-prompt perceptual spread under this protocol.

D.6 Budget and Stage Diagnostic Protocols

Candidate construction. Both budget experiments use SD3.5M with 40 confirmation prompts, 1024×1024 images, 40 sampling steps, CFG 7, and seed 42. At $p = 36$, positive-gradient candidates are completed and selected by strict target improvement, with baseline rollback. The experiments differ in their saved candidate pools: Figure 4(c) uses radii $(0.032, 0.008)$, while Table 21 uses a separately generated pool with menu

$$\boldsymbol{a} = (0.032, 0.008, 0.016, 0.064, 0.004, 0.128, 0.002, 0.256).$$

Within the latter pool, budget $K \in \{1, 2, 4, 8\}$ uses the first K completed candidates, excluding the baseline. HPSv2 and Aesthetic are used for reporting rather than proposal construction or selection.

Paired statistics. For baseline x_i^0 and output $x_{i,s,K}$ returned by selector s, the gain under reporting evaluator m is

$$\overline{\Delta}_{m;s,K} = \frac{1}{40} \sum_{i=1}^{40} \big[m(x_{i,s,K}) - m(x_i^0) \big].$$

Thus, target gain sets $m = s$, while transfer uses a separate reporting evaluator. GenEval diagnostics use the dense internal score rather than the official four-seed aggregate. Acceptance is the fraction of prompts returning a non-baseline candidate; rollback contributes zero gain. Confidence intervals are obtained by prompt-level bootstrapping.

Scorer-selected stage responses. Figure 6 uses positive-gradient proposals, two radii, same-evaluator selection, and rollback at $p \in \{1, 8, 24, 36\}$. For each position, target gain, low-frequency RMS, high-frequency energy share, and CLIP-image semantic displacement are averaged over all 40 prompts; rollback contributes zero. These means also define the stage comparisons: for mean target gain $g_s(p)$, $g_s(1)/g_s(36)$ gives the dense-GenEval early advantage, while $g_s(36)/g_s(1)$ gives the PickScore and CLIPScore late advantages, yielding $3.7\times$, $1.9\times$, and $2.6\times$.

Random-probe reachability. To characterize the suffix independently of target selection, we evaluate four signed random perturbations per prompt and position at relative radius 0.032. Averaging over perturbations and prompts gives low-frequency RMS $L(p)$ and high-frequency energy share. The endpoint values are reported in Table 22.

Table 22: **Random-probe endpoint statistics.**

Diagnostic	$p = 1$	$p = 36$
Low-frequency RMS	0.075794	0.004157
Low-frequency RMS / value at $p = 1$	1.000	0.055
High-frequency energy share	0.461	0.824

These endpoints and the early/late position groups define two distinct ratios: the group contraction and the normalized endpoint response are, respectively,

$$\frac{[L(1) + L(8)]/2}{[L(24) + L(36)]/2} = 10.428451 \simeq 10.43, \qquad \frac{L(36)}{L(1)} \simeq 0.054842 \simeq 0.055.$$

E Mathematical Details

E.1 Discrete Adjoint for the Implemented Sampler

Let $h = z^{(p)}$ and let $z^{(i+1)} = F_i(z^{(i)}; y)$ for $i = p, \ldots, N-1$. For a differentiable terminal objective

$$\mathcal{J}_p(h; y) = J_\phi(D(z^{(N)}), y), \tag{16}$$

define the terminal adjoint

$$\lambda_N = \nabla_{z^{(N)}} J_\phi(D(z^{(N)}), y). \tag{17}$$

Applying the chain rule backward through each implemented transition gives

$$\lambda_i = \Big(DF_i(z^{(i)}; y) \Big)^\top \lambda_{i+1}, \qquad i = N-1, \ldots, p. \tag{18}$$

Therefore

$$\lambda_p = \nabla_h J_\phi(D(G_{p:N}(h; y)), y)\,, \tag{19}$$

which is the reverse-mode derivative of the actual discrete suffix. The sampler, guidance configuration, time grid, numerical solver, and decoder in this computation are identical to those used for candidate generation.

E.2 SELECTOR-RELATIVE MONOTONICITY

Proposition E.1. *Let h be the incoming state, let $\mathcal{C}_p(h)$ be the finite candidate set, and assume the suffix and terminal selector are deterministic. If CLOSER returns the highest-scoring admissible element of $\{h\} \cup \mathcal{C}_p(h)$, then*

$$V(x(h^\star), y) \geq V(x(h), y). \tag{20}$$

Proof. The incoming state h is an available fallback. A highest-scoring admissible choice therefore cannot have a lower selector score than h: it either retains the incoming score, including in a tie, or returns a higher-scoring candidate. Both cases give the stated inequality. □

This guarantee concerns the configured selector and admissibility rule. Transfer to other evaluators is therefore assessed separately in Figure 4(b) and Appendix D.4, using the evaluator roles specified in Table 10.

E.3 PER-STAGE OPERATION COUNT

At stage r, let K_r be the correction-iteration budget, $|\mathcal{D}_r|$ the number of proposal directions, and $|\mathcal{A}_r|$ the number of displacements. Denoting one differentiable suffix evaluation by $C_{\mathrm{grad}}(p_r)$ and one completed candidate suffix by $C_{\mathrm{fwd}}(p_r)$, the worst-case stage operation count is

$$K_r \left[C_{\mathrm{grad}}(p_r) + |\mathcal{D}_r|\,|\mathcal{A}_r|\, C_{\mathrm{fwd}}(p_r) \right]. \tag{21}$$

This expression separates gradient computation from completed candidate rollouts. Both terms depend on the remaining suffix length, while the finite menus bound the number of rollouts. Empirical comparisons measure the resulting work through the normalized end-to-end wall-clock metric in Section 4.1.

E.4 SEQUENTIAL CORRECTION

For ordered positions $p_1 \leq \cdots \leq p_R$, the accepted trajectory from one stage supplies the starting point for the next. After an accepted intervention, CLOSER recomputes the terminal gradient at the updated state; between control positions, it executes the unchanged frozen sampler. This lets the one-stage selection rule be applied repeatedly.

Sequential consequence. Under a common deterministic selector V, retaining the incoming state at every stage extends Proposition E.1 across ordered interventions: $V(x^{(R)}, y) \geq \cdots \geq V(x^{(1)}, y) \geq V(x^{(0)}, y)$. Applying the proposition at each stage and chaining the resulting inequalities gives the sequence above.

Guarded proposal–selector consequence. For auxiliary terminal scores P_j and tolerances $\delta_j \geq 0$, let $\mathcal{A}_p(h)$ contain the incoming state and every completed candidate satisfying $P_j(x(h'), y) \geq P_j(x(h), y) - \delta_j$ for all j. Selecting $h^* \in \arg\max_{h' \in \mathcal{A}_p(h)} V(x(h'), y)$ guarantees both $V(x(h^*), y) \geq V(x(h), y)$ and every registered auxiliary floor. These floors are relative to the incoming image at the current correction. Because the invariant depends on terminal selection rather than proposal construction, J may differ from V, and V need not be differentiable. The same selection principle thus covers the decoupled selectors in Table 20 and the guarded policies in Table 13.