Training-Free Adaptive Sampling Scheduler for Discrete-Time Consistency Distillation

One scheduler, no retraining — drop it into LCM, PCM, TCD or TDD.

Qi Wang✉ · **Jinjia Zhou**

Graduate School of Science and Engineering · Hosei University · Tokyo, Japan

Scan: full summary

1 Why few-step sampling still breaks

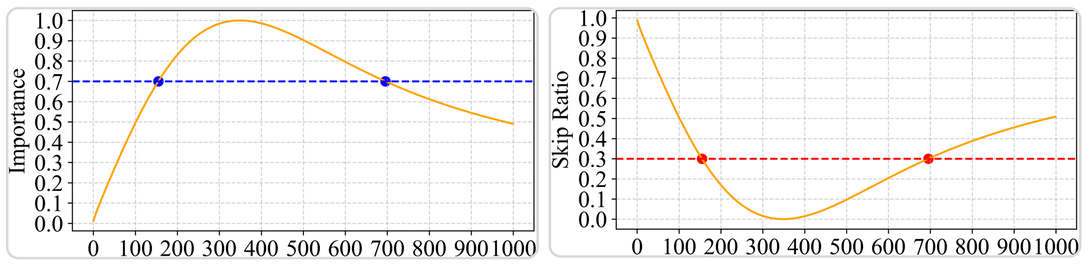
- **Scheduler rigidity.** Every discrete-time method ships a schedule welded to its own framework and step budget; deterministic- (LCM, PCM) and stochastic-target (TCD, TDD) strategies alike degrade as soon as the scheduler or the step count changes.
- **CFG compounding.** Inference guidance multiplies with the inner distillation scale, pushing pixels into saturation and colour drift.

2 Timesteps are not equally informative

In the forward process the image stays legible to $t \approx 200$ and is indistinguishable from noise past $t \approx 800$; nearly all structural change is packed in between. We quantify this with an **Importance** I_t — the rate of change of log-SNR, inverted and normalised to $[0, 1]$. A large I_t marks a timestep where SNR moves fast; the complementary **skip rate** R_t says how much of that step can safely be skipped.

IMPORTANCE & SKIP RATE (EQ. 8) | $\mathrm{SNR}_t = \bar{\alpha}_t/(1-\bar{\alpha}_t)$, $\varepsilon = 10^{-8}$

$$I_t = \frac{\left|\nabla_t \ln(\mathrm{SNR}_t + \varepsilon)\right|^{-1}}{\max_{0 \le j < T} \left|\nabla_j \ln(\mathrm{SNR}_j + \varepsilon)\right|^{-1}}, \qquad R_t = 1 - I_t + \varepsilon$$



(a) I_t peaks near $t \approx 350$; the dashed line is $\theta = 0.7$, whose two crossings bracket the band worth sampling densely. **(b)** The skip rate $R_t = 1 - I_t$ mirrors it.

3 Adaptive schedule & γ-I sampler ★ KEY

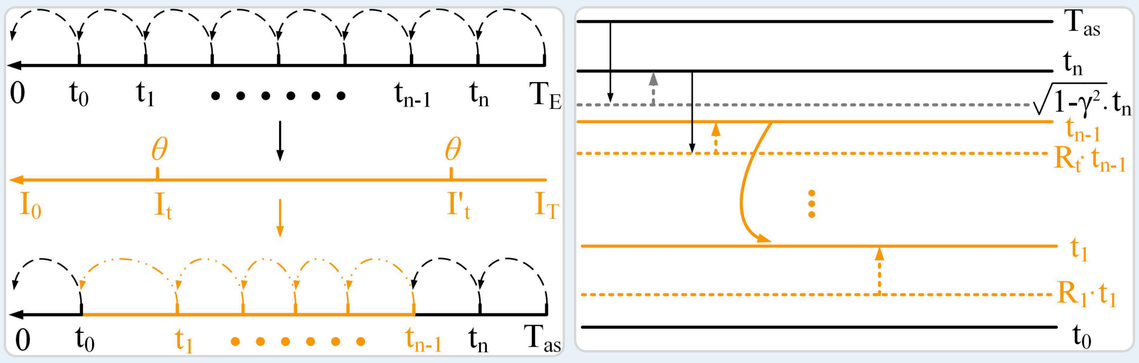
The schedule: keep uniform-grid timesteps T_E where change is slow, substitute high-Importance candidates T_I where it is fast. The **step count never changes**, only its spacing. **The sampler:** instead of CTM's fixed γ, each step scales its noise re-injection by the local Importance, widening exploration where the trajectory turns fastest.

ADAPTIVE SCHEDULE (EQ. 9), $\theta = 0.7$

$$T_{\mathrm{AS}} = \{\, t_i \in T_I \mid I_{t_i} > \theta \,\} \cup \{\, t_i \in T_E \mid I_{t_i} \le \theta \,\}$$

UNIFORM STEP (EQ. 10) / HIGH-IMPORTANCE STEP (EQ. 11)

$$x_{t_{n+1}} \xrightarrow{f_\theta} \sqrt{(1-\gamma)^2}\cdot t_n \to x_{t_n}, \qquad x_{t_n} \xrightarrow{f_\theta} I_t \cdot t_{n-1} \to x_{t_{n-1}}$$



(a) Dense sampling in high-Importance regions, sparse elsewhere, at a preserved budget. **(b)** Importance-modulated forward and backward transitions lift sampling quality precisely where the SNR changes most rapidly.

4 Smoothing & colour balancing

To undo CFG compounding we squash outliers with $\tanh$, then re-centre channel and global means so values return to its **linear region** instead of pinning at ± 1 — no prior deformation, and more information survives under strong guidance.

SOFT CLAMP & RE-CENTRING (EQ. 12–13), $\alpha = \beta = 0.5$

$$x_0 = \frac{e^{x_0} - e^{-x_0}}{e^{x_0} + e^{-x_0}}, \qquad x_c \leftarrow x_c - \alpha\,\mathrm{mean}(x_c), \qquad x_0 \leftarrow x_0 - \beta\,\mathrm{mean}(x_0)$$

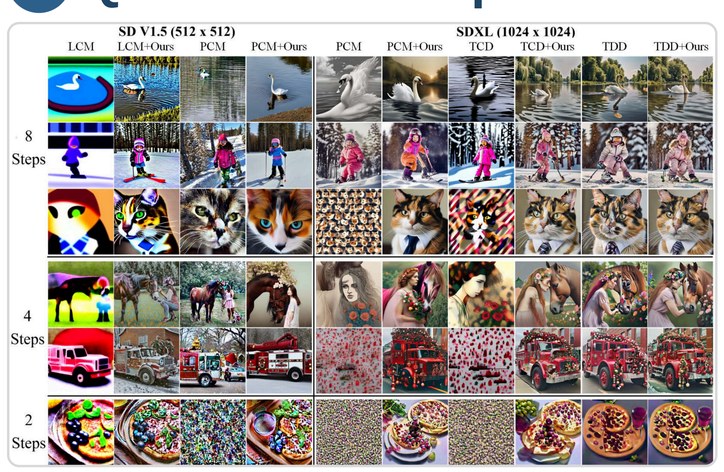
5 Quantitative results ★ Headline

MS-COCO 2014 val · Karpathy 30k captions · CFG $\omega = 7.5$ · columns 2 / 4 / 8 steps · △ MPI = mean gain.

Method	FID ↓			CLIP Score ↑			Inception Score ↑		
	2	4	8	2	4	8	2	4	8
SDXL — 1024×1024									
PCM	372.82	112.65	31.73	18.44	24.18	30.44	1.71	11.67	25.78
PCM + Ours	**65.83**	**29.40**	**23.21**	**30.22**	**30.40**	**31.52**	**16.54**	**24.59**	**31.80**
TCD	363.50	103.66	53.72	18.73	26.05	30.44	1.82	12.47	17.83
TCD + Ours	**62.89**	**28.51**	**27.44**	**28.88**	**31.71**	**32.06**	**16.33**	**28.02**	**32.17**
TDD	58.45	29.80	27.72	29.27	31.12	31.47	17.18	27.02	29.72
TDD + Ours	**55.71**	**27.88**	**26.60**	**29.49**	**31.47**	**31.74**	**17.62**	**29.36**	**32.88**
△ MPI	**203.45**	**53.44**	**11.97**	**7.38**	**4.08**	**0.99**	**9.93**	**10.27**	**7.84**
SD v1.5 — 512×512									
LCM	86.33	88.20	109.84	28.02	26.48	25.19	14.28	11.73	8.71
LCM + Ours	**58.01**	**30.04**	**47.67**	**30.18**	**30.71**	**30.79**	**17.18**	**28.81**	**19.84**
PCM	424.04	89.99	38.82	18.99	26.51	30.02	1.76	12.49	21.07
PCM + Ours	**60.66**	**23.11**	**22.47**	**29.93**	**30.37**	**31.03**	**16.86**	**28.93**	**31.65**
△ MPI	**195.85**	**62.52**	**39.26**	**6.55**	**4.05**	**3.31**	**9.00**	**16.76**	**10.86**

Reading the table. All five baselines improve on every metric at every budget, and the gains land where few-step sampling is weakest — at 2 steps FID falls **82–86%** for PCM and TCD. LCM alone gets *worse* with more steps (86.33 → 109.84 FID): the CFG-overexposure signature, which ours reverses. The pattern holds across both resolutions and all four distilled students, not just on the metric we optimised for.

6 Qualitative comparison



What to look for

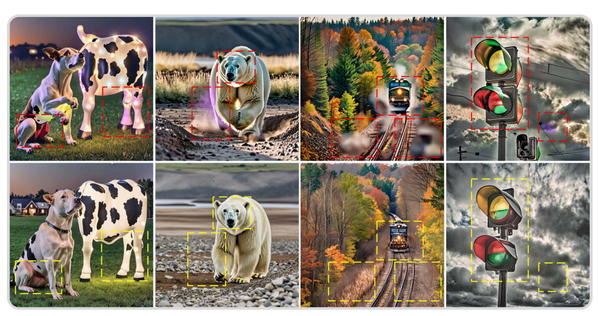
At **2 and 4 steps** the baselines lose global structure and wash the colour out. Ours keeps the layout coherent and the detail sharper — on **both** backbones. By 8 steps the gap narrows, as expected once the student has budget enough to correct itself.

7 Ablations



The sampler

Scaling noise re-injection by the local Importance, instead of CTM's fixed γ, gives more stable shapes and fewer structural errors at the same step budget. Threshold sweep: best $\theta \in [0.7, 0.8]$.



Smoothing & colour balance

Top row: the clipping used by prior work. **Bottom row:** ours — natural colour, fewer saturated regions and more surviving detail at high CFG.

8 What to take away ★

- **Diagnosis.** The schedule, not the student, is the bottleneck under cross-method or variable-step transfer.
- **Mechanism.** An SNR-derived Importance says where computation is worth spending; selection, the γ-I sampler and a colour pass do the rest.
- **Cost.** No retraining, no extra parameters, negligible runtime overhead — a drop-in scheduler swap on top of an existing distilled student.

In short: training-free, lightweight, drop-in — largest gains at 2–4 steps, consistent across SDXL and SD v1.5, across all four distilled students tested.

Adaptive Sampling Scheduler · ICANN 2026 · Supported by JSPS KAKENHI JP25K15165 and the Kayamori Foundation of Informational Science Advancement.

btaeyang.github.io/icann2026 · qi.wang.7c@stu.hosei.ac.jp