On the Collapse of Generative Paths: A Criterion and Correction for Diffusion Steering

Presenter: **Ziseok Lee**

Project Page

- Naively composing pretrained generative experts can create invalid intermediate probability paths.
- ACE diagnoses and repairs this before sampling.

Why Compose Generative Experts?

- Many tasks require several constraints at once:

$$p(X, Y \mid A, B) \propto p(X \mid A) \frac{p(X, Y \mid B)}{p(X)}$$

- Local expert $p(X \mid A)$ & global expert $p(X, Y \mid B)$ & marginal correction $p(X)$
  = multi-constraint target

Task	Local constraint	Global constraint
Synthetic checker	$(X \geq 0)$	$(X + Y \geq 0)$
Scaffold decoration	scaffold topology	protein-pocket binding
Multi-instance images	regional object prompts	global image prompt

Hidden Failure: The Sampler Runs, but the Path Does Not Exist

- Define the ratio-of-densities:

$$h_t(x) = \prod_i \left(\tilde{q}_t^{(i)}(x)\right)^{\gamma_i(t)}$$

- h_t defines a probability path:

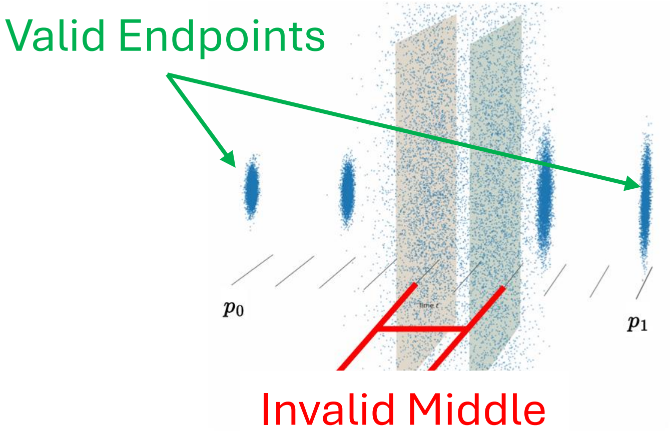
$$p_t^*(x) = \frac{h_t(x)}{Z_t}, \qquad Z_t = \int h_t(x)dx$$

- However,

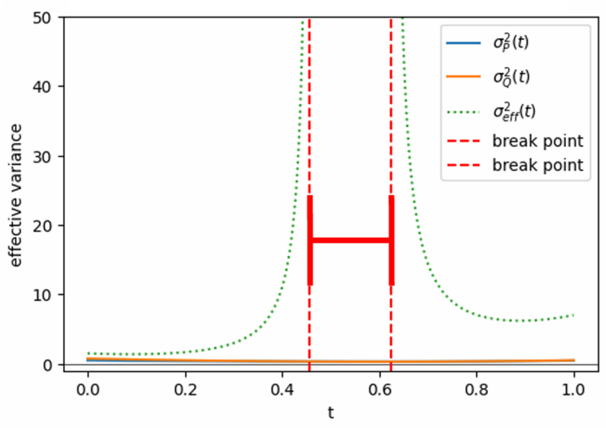
$$\boxed{Z_t = \infty \quad \Rightarrow \quad p_t^* \text{ does not exist}}$$

- (Definition) ***Marginal Path Collapse*** occurs when an intermediate composed density becomes non-normalizable and ceases to exist despite valid endpoints.
- Numerical solvers (e.g., SDE, ODE) still run but follow an unintended path and diverge from the true target.

(a) Non-integrable Region in Gaussian Path



(b) Effective Variance across time



Path Existence Criterion

- How can we detect path collapse?
- For each coordinate $k \in \{1, \dots, d\}$, and time $t \in [0,1]$, define the criterion:

$$C_k(t) = \sum_{i:k \in I_i} \frac{\gamma_i(t)}{\left(\alpha_t^{(i)}\right)^2}, \qquad C(t) = \min_k C_k(t)$$

where I_i is the set of coordinates that expert i acts on, and $\gamma_i(t), \alpha_t^{(i)}$ are the exponents and the noise schedule for expert i, respectively.

- (Theorem) Sufficient Criterion for Path Existence and Path Collapse

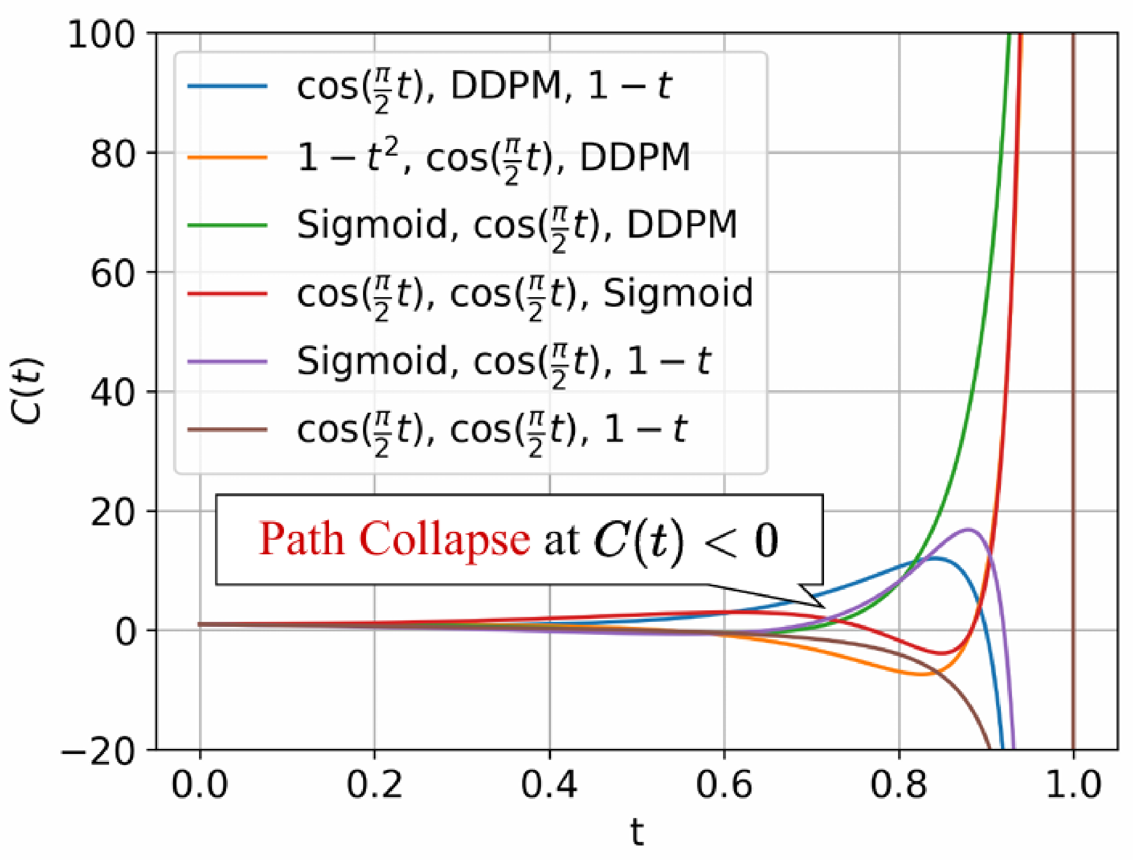
$$\begin{aligned} C(t) > 0 \quad &\Rightarrow \quad Z_t < \infty \\ C(t^*) < 0 \quad &\Rightarrow \quad Z_{t^*} = \infty \end{aligned}$$

- Boundary note: When $C(t) = 0$, either outcome is possible. See Example B.1 in the Appendix.

Common Noise Schedules and Marginal Path Collapse

- Many heterogeneous three-expert compositions enter a region where $C(t) < 0$, implying non-normalizable intermediate densities.

(a) Evolution of Criterion $C(t)$ w/o ACE



ACE: **A**daptive Path **C**orrection with **E**xponents

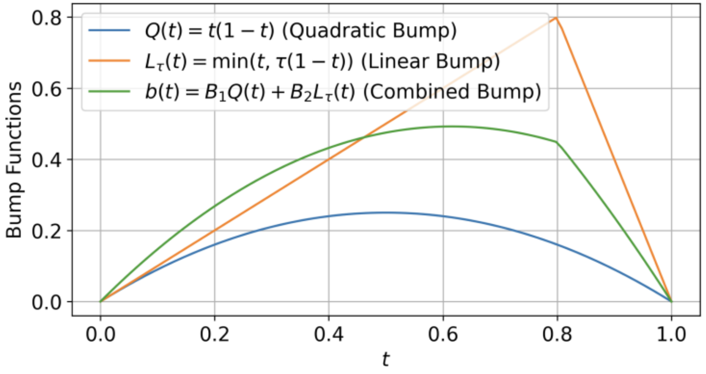
- Three-step pipeline:

1. **Diagnose** $C(t)$
2. **Correct** exponents with endpoint-preserving bump functions $b_i(t)$
   $\tilde{\gamma}_i(t) = \gamma_i(t) + b_i(t)$ where $b_i(0) = b_i(t_{\text{end}}) = 0$ to ensure $C(t) \geq \delta > 0$ for all t
3. **Sample** the corrected path with weighted particles and sequential Monte Carlo

- Such bump functions $b_i(t)$ are not unique. We provide a small concrete family:

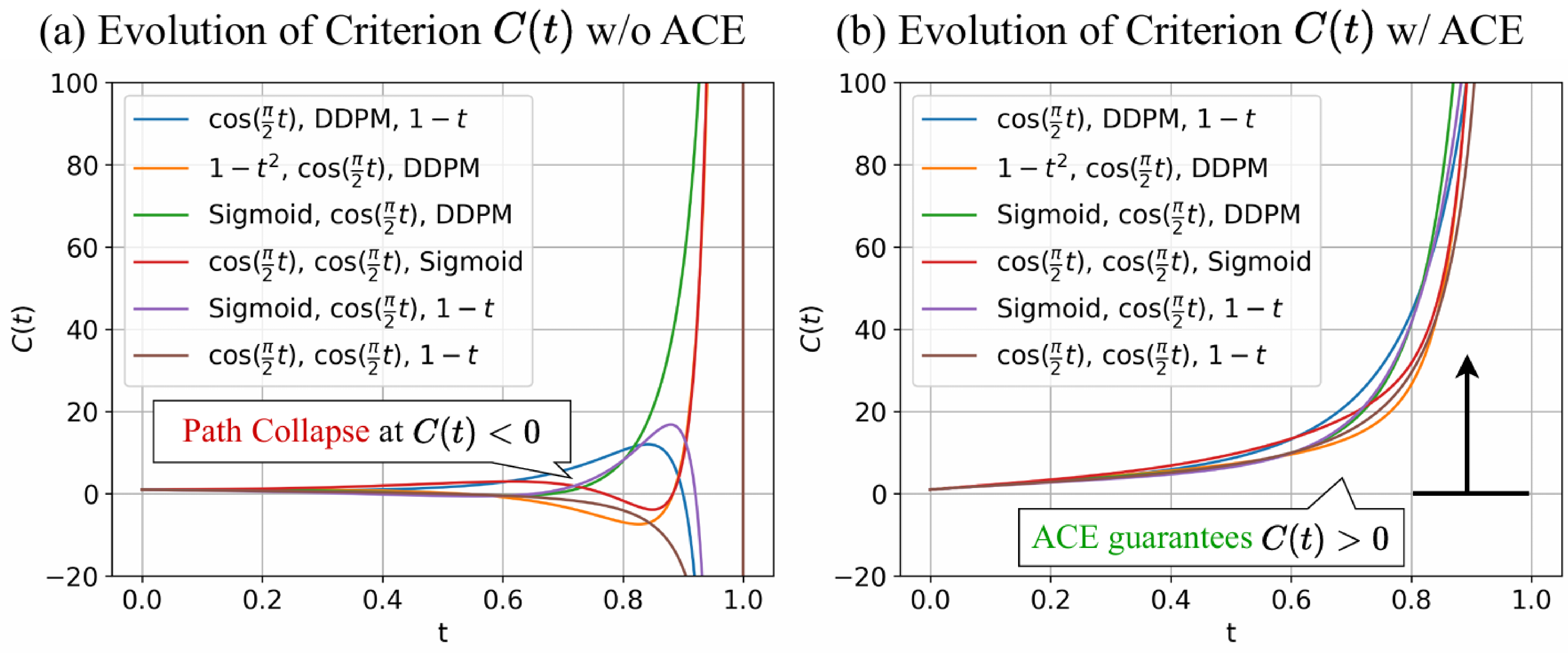
$$b(t) = B_1 t(1-t) + B_2 \min\big(t, \tau(1-t)\big)$$

- We prove that there always exists coefficients B_1, B_2, τ that guarantee $C(t) > 0$ for all t.



Common Noise Schedules and ACE

- The corrected exponents of ACE ensure $C(t) > 0$ for all sampled times:



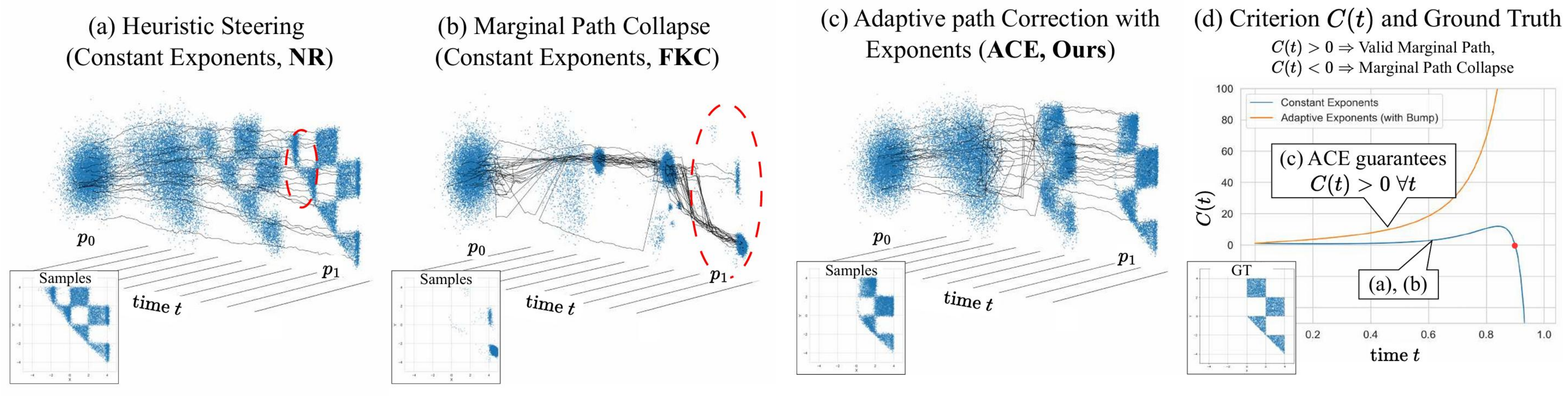
Synthetic Checker: Collapse is Visible

- Multi-Constraint Target Distribution

$$p(\,(x,y)\mid x\geq 0, x+y\geq 0\,)\propto p(\,x\mid x\geq 0\,)\,\frac{p(\,(x,y)\mid x+y\geq 0\,)}{p(x)}$$

Here, $(x,y)\in[-4,4]^2$ are compactly supported and coupled via the checkerboard distribution:

$$p_{x,y}(x,y)=\frac{1}{32}\cdot\mathbf{1}\left[\left\lfloor\frac{x}{2}\right\rfloor+\left\lfloor\frac{y}{2}\right\rfloor \text{ is an odd integer}\right]\cdot\mathbf{1}[(x,y)\in[-4,4]^2]$$



Synthetic Checker: Collapse is Visible

- Quantitative Results (average of 5 seeds):

Method	W_1 (↓)	W_2 (↓)	MMD (↓)
NR	0.89	1.18	0.092
FKC	1.37	1.59	0.419
ACE (Ours)	**0.20**	**0.29**	**0.012**

  - ACE reduces distribution error by over 4 ×.

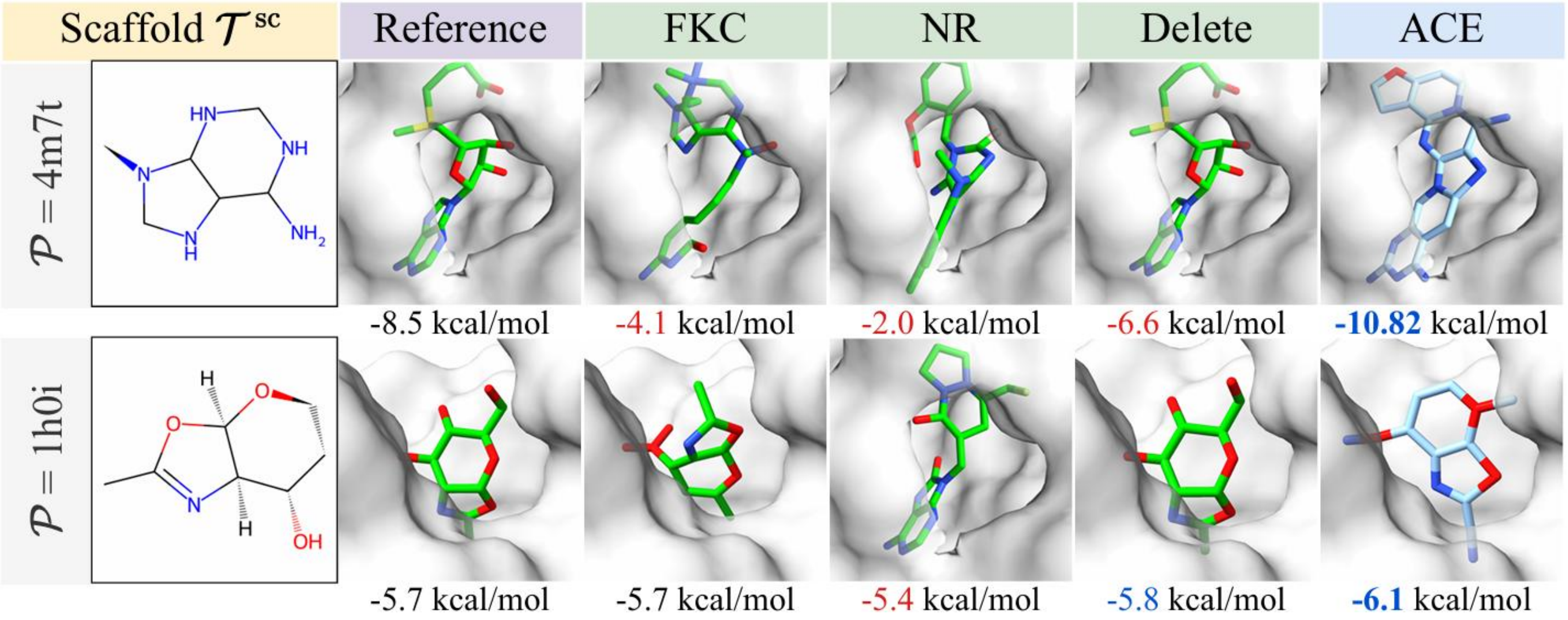
- NR (No Resampling) and FKC (Feynman-Kac Correctors) fail under the collapsed heterogeneous schedule.
- ACE with $B_1 = 10, B_2 = 1.5$ gives the best distributional metrics as shown above.

Flexible-Pose Scaffold Decoration

- Multi-Constraint Target Distribution:

$$p(M^{\mathrm{sc}}, M^{\mathrm{R}} \mid T^{\mathrm{sc}}, P) \propto p(M^{\mathrm{sc}} \mid T^{\mathrm{sc}}) \frac{p(M^{\mathrm{sc}}, M^{\mathrm{R}} \mid P)}{p(M^{\mathrm{sc}})}$$

Meaning: Generate a molecule $M = (M^{sc}, M^{R})$ composed of a scaffold part and the remainder with two constraints: (1) M^{sc} preserves a given topology T^{sc} and (2) the molecule binds to a given protein pocket P.



- Baselines: FKC, NR (multi-expert steering methods), and Delete (a task-specific drug optimization model)

Flexible-Pose Scaffold Decoration

- Quantitative Results (Evaluated on the CrossDocked Test Set)

Method	PEC	OSR ↑	Avg Vina ↓	Avg QED ↑
FKC	✗	0.40	-6.24	0.44
NR	✗	0.47	-6.35	0.40
Delete	-	0.47	-5.07	**0.55**
ACE-lite (Ours)	✓	<u>0.66</u>	<u>-7.00</u>	0.50
ACE (Ours)	✓	**0.75**	**-7.10**	<u>0.53</u>

- Computational Cost

Method	ACE	ACE-lite	FKC/NR	Delete
Time (min/molecule)	0.34	**0.19**	**0.19**	0.31
VRAM (GB)	1.57	1.54	1.53/1.54	**1.46**

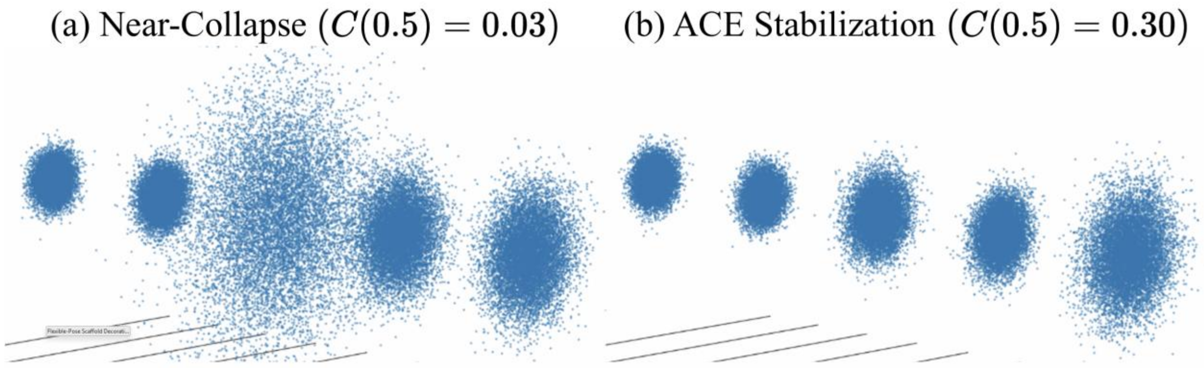
Broader Results

- On multi-instance image generation (COCO-MIG), ACE improves attribute success by +9.57%p.

Method	Backbone	L2	L3	L4	L5	L6	Avg	CLIP↑	Local CLIP↑
Base Diffusion Model with global prompt									
SD1.5	SD1.5	5.59 (18.83)	4.79 (17.43)	2.83 (14.95)	2.41 (13.93)	2.21 (15.94)	3.11 (15.75)	24.64	18.36
SDXL	SDXL	5.59 (19.78)	4.48 (18.54)	2.81 (16.67)	2.14 (15.72)	2.80 (18.42)	3.17 (17.55)	25.71	18.63
SD3.5-M	SD3.5-M	8.05 (21.57)	8.46 (21.37)	6.07 (18.98)	5.02 (17.39)	4.40 (17.80)	5.86 (18.85)	26.41	18.77
Flux.1-dev	Flux.1-dev	8.83 (22.00)	7.50 (20.93)	4.86 (17.75)	4.53 (16.77)	3.22 (16.49)	5.08 (18.03)	26.17	18.56
*Inference-time Control using **only** the base Diffusion Model*									
NR	SD1.5	24.61 (30.50)	24.19 (29.81)	19.36 (25.64)	17.67 (24.42)	19.54 (25.91)	20.24 (26.53)	24.96	20.21
FKC	SD1.5	33.02 (34.43)	31.18 (33.34)	28.80 (31.53)	25.46 (29.63)	22.88 (28.61)	28.27 (31.51)	25.44	20.66
ACE ($B_1 = 5, B_2 = 0$)	SD1.5	45.31 (41.48)	42.50 (40.60)	36.25 (35.20)	32.75 (32.66)	32.40 (33.40)	37.84 (36.67)	25.59	21.18
NR	SD2.1	26.52 (31.18)	25.76 (30.50)	20.94 (27.24)	18.83 (24.77)	18.78 (25.31)	21.04 (26.93)	24.89	19.96
FKC	SD2.1	39.69 (37.78)	34.93 (35.91)	31.98 (33.95)	28.42 (32.07)	24.93 (30.02)	31.99 (33.95)	25.27	20.64
ACE ($B_1 = 5, B_2 = 0$)	SD2.1	46.25 (42.24)	41.04 (39.45)	41.72 (37.96)	38.38 (37.17)	34.90 (34.82)	40.46 (38.33)	24.94	21.09
Adapter rendering methods (Requires Additional Training on new data or External Models)									
GLIGEN	SD1.4	38.36 (33.96)	32.79 (29.58)	28.67 (25.95)	25.02 (23.88)	26.98 (24.93)	28.84 (26.47)	24.91	20.78

"A photo of a brown boat and a brown boat and a yellow boat and a black boat and a white boat and a green boat"

- Even when path collapse does not occur, ACE stabilizes sampling and yields additional gains.



Takeaways

- The scope:

ACE is useful when a target can be written as a **ratio or product of existing expert densities** in a common sampling space.

- The takeaway is:

diagnose with PEC, **repair** with ACE, and then **sample** the corrected path.

Project Page

Code

Paper