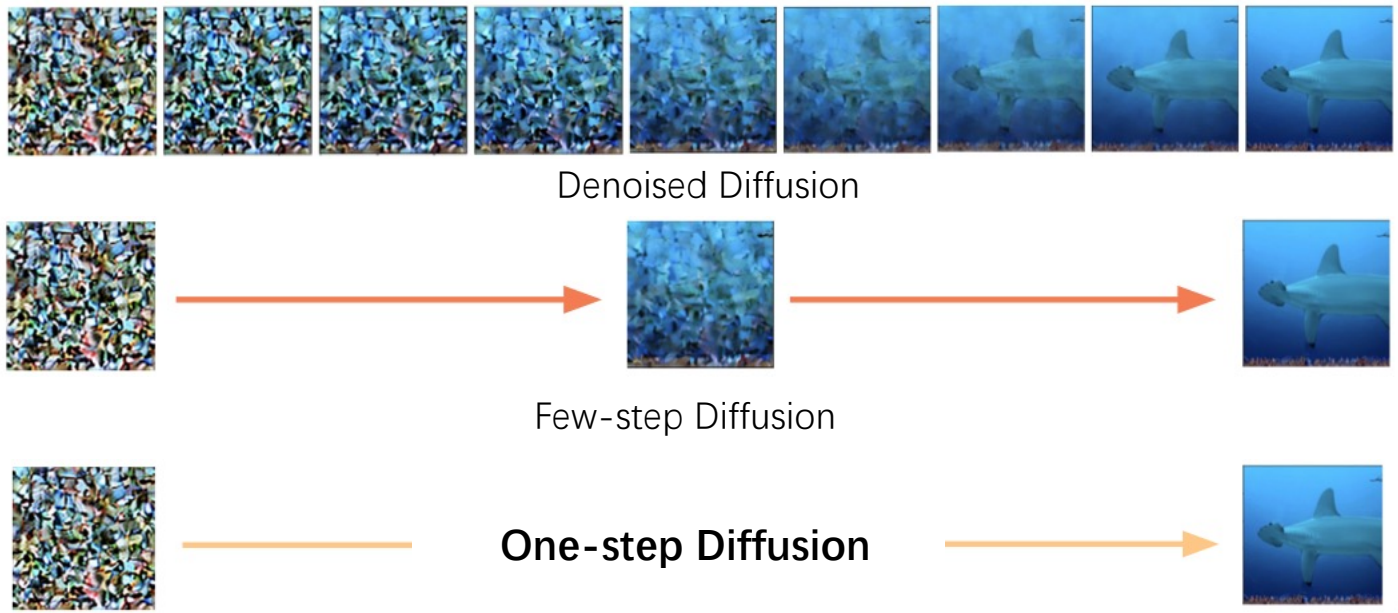
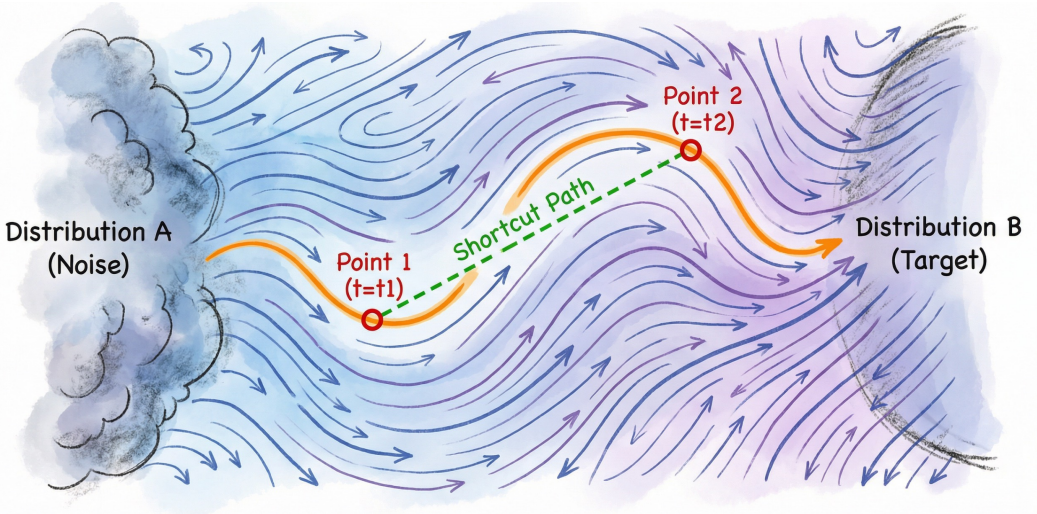
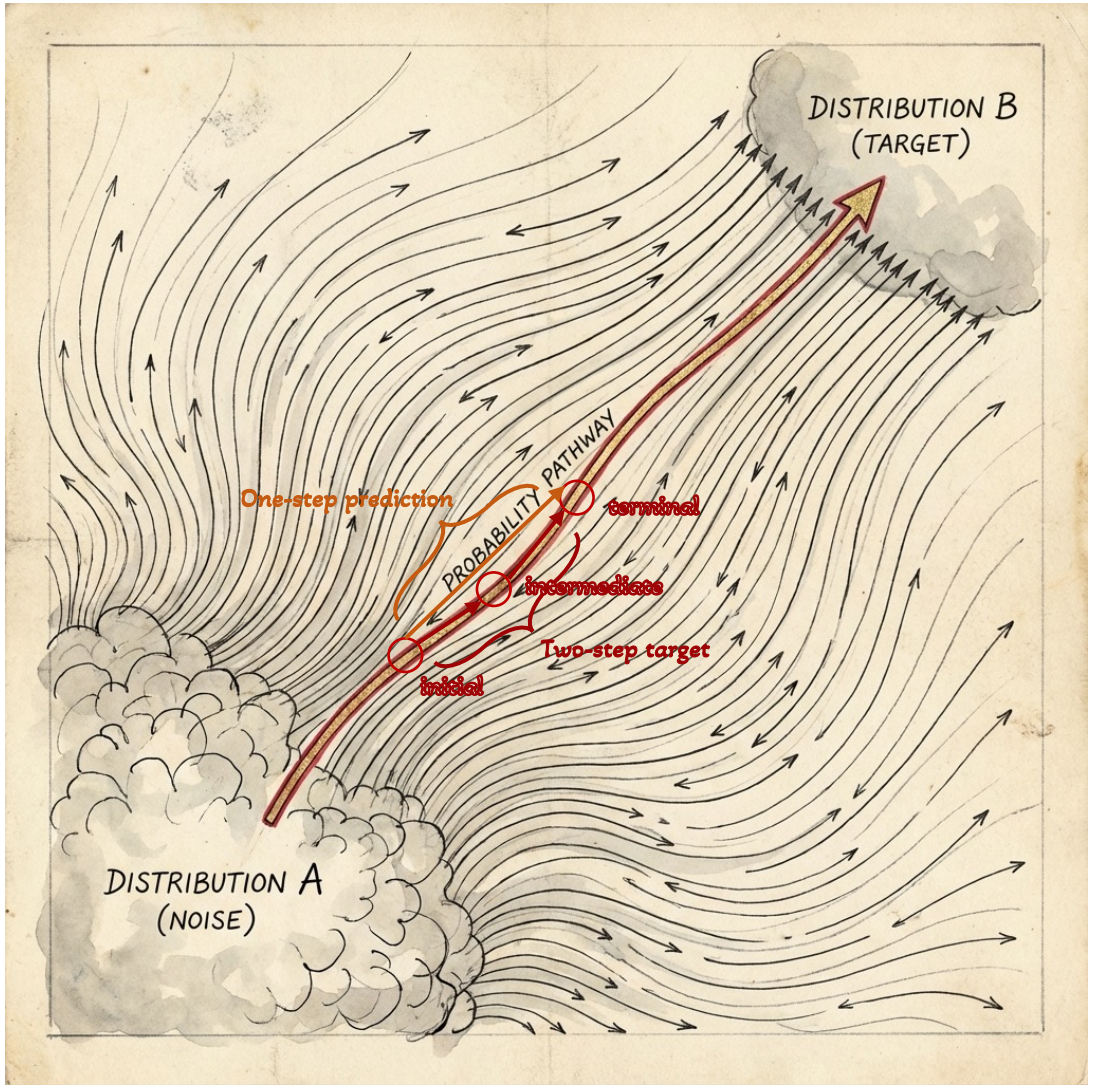
On the design space of one step diffusion via shortcutting flow paths

Reporter: Haitao Lin *Supervisor: Stan Z. Li & Tailin Wu*

School of Engineering, Westlake University
Computer Science and Technology, Zhejiang University






DISTRIBUTION B
(TARGET)
PROBABILITY PATHWAY
One-step prediction
terminal
intermediate
Two-step target
initial
DISTRIBUTION A
(NOISE)

Preliminary

1. Diffusion Models

$$\boldsymbol{x}_t = \alpha_t \boldsymbol{x}_0 + \sigma_t \boldsymbol{\varepsilon} \qquad \boldsymbol{x}_0 \sim p_0, \boldsymbol{\varepsilon} \sim p_1$$

(Diffusion process)

$$\boldsymbol{v}_t(\boldsymbol{x}) = \dot{\alpha}_t \mathbb{E}(\boldsymbol{x}_0 | \boldsymbol{x}_t = \boldsymbol{x}) + \dot{\sigma}_t \mathbb{E}(\boldsymbol{\varepsilon} | \boldsymbol{x}_t = \boldsymbol{x})$$

(Marginal Velocity)

$$\dot{\boldsymbol{x}}_t = \boldsymbol{v}_t(\boldsymbol{x}_t)$$

(PF-ODE)

$$\boldsymbol{v}_{t|0} = \boldsymbol{v}_t(\boldsymbol{x}_t | \boldsymbol{x}_0) = \dot{\alpha}_t \boldsymbol{x}_0 + \dot{\sigma}_t \boldsymbol{\varepsilon}$$

(Conditional Velocity)

$\sigma = \sigma_{\text{data}}$, $\alpha_t = \cos(\frac{\pi}{2}t)$, and $\sigma_t = \sin(\frac{\pi}{2}t)$

(Coefficient in cosine paths)

$\sigma = 1$, $\alpha_t = 1 - t$ and $\sigma_t = t$

(Coefficient in linear paths)

Preliminary

2. Flow Map Solvers

$$\boldsymbol{x}_r = X_{t,r}(\boldsymbol{x}_t) = \boldsymbol{x}_t + \int_t^r \boldsymbol{v}_\tau(\boldsymbol{x}_\tau)d\tau,$$

(Flow Map from t to r)

$$X_{t,r}(\boldsymbol{x}_t) = \boldsymbol{x}_t + (r-t)\cdot \boldsymbol{u}_{t,r}(\boldsymbol{x}_t)$$

$$\text{where} \quad \boldsymbol{u}_{t,r}(\boldsymbol{x}_t) = \frac{1}{r-t}\int_t^r \boldsymbol{v}_\tau(\boldsymbol{x}_\tau)\,d\tau,$$

(Average Velocity Solver)

$$X_{t,r}(\boldsymbol{x}_t) \approx \mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_t, t, r) = \bar{\alpha}_{t,r}\boldsymbol{x}_t + \bar{\beta}_{t,r}\boldsymbol{v}_t,$$

(One-step DDIM Solver)

$$\bar{\alpha}(t,r) = \cos\left(\frac{\pi}{2}(r-t)\right), \quad \bar{\beta}(t,r) = \frac{2}{\pi}\sin\left(\frac{\pi}{2}(r-t)\right)$$

(Coefficient in cosine paths)

$$\bar{\alpha}(t,r) = 1, \quad \bar{\beta}(t,r) = r-t$$

(Coefficient in linear paths)

General Designing Framework

1. Consistency Property & Learning Objective & Time Sampler

$$X_{s,r}(X_{t,s}(\boldsymbol{x}_t)) = X_{t,r}(\boldsymbol{x}_t).$$ (Consistency Property)

$$\arg\min_{\theta} \mathbb{E}_{r,s,t\sim p(\tau),\, \boldsymbol{x}_t\sim p_t}\Big[\underbrace{w(r,s,t)\cdot d\big(\overbrace{X^{\theta}_{t,r}(\boldsymbol{x}_t)}^{\text{one-step prediction}},\ \overbrace{\mathrm{sg}(\hat{X}_{s,r}\circ\hat{X}_{t,s}(\boldsymbol{x}_t))}^{\text{two-step target}}\big)}_{l(\boldsymbol{x}_t,r,s,t;\theta)}\Big]$$ (Learning Objective)

where w is the weight term, $\hat{X}$ and X^{θ} are flow maps obtained with the conditional velocity or the neural network F^{θ}, $d(\cdot,\cdot)$ is a loss metric function, such as the squared l_2-distance, and $sg(\cdot)$ is the stop gradient operator in backpropagation.

We call $\hat{X}_{s,r}\circ\hat{X}_{t,s}(x_t)$ two-step flow map targets and $X^{\theta}_{t,r}(x_t)$ one-step flow map predictions, and write the inner loss term of expectation as $l(x_t, r, s, t; \theta)$.

$\{r,s,t\}$ discrete time points ⟹ DTSC $\quad s\to t$ ⟹ CTSC (Time Sampler)

Examples: DTSC & CTSC

1. Consistency Training Model (DTSC)

$$
\begin{aligned}
&\boldsymbol{x}_0 \sim p_0, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0,1) \\
&\boldsymbol{x}_t = \cos(\frac{\pi}{2}t)\boldsymbol{x}_0 + \sin(\frac{\pi}{2}t)\boldsymbol{\varepsilon} \\
&\boldsymbol{v}_{t|0} = -\frac{\pi}{2}\sin(\frac{\pi}{2}t)\boldsymbol{x}_0 + \frac{\pi}{2}\cos(\frac{\pi}{2}t)\boldsymbol{\varepsilon} \\
&\hat{\boldsymbol{x}}_s = \cos(\frac{\pi}{2}s)\boldsymbol{x}_0 + \sin(\frac{\pi}{2}s)\boldsymbol{\varepsilon},
\end{aligned}
$$

(Conditional Flow Path)

$$
\begin{aligned}
\hat{X}_{t,s}(\boldsymbol{x}_t) &= \hat{\boldsymbol{x}}_s = \mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_{t|0}, t, s) \\
\hat{X}_{s,r}(\hat{\boldsymbol{x}}_s) &= \hat{\boldsymbol{x}}_r = \mathrm{DDIM}(\hat{\boldsymbol{x}}_s, F^\theta(\hat{\boldsymbol{x}}_s), s, r) \\
X^\theta_{t,r}(\boldsymbol{x}_t) &= \boldsymbol{x}^\theta_r = \mathrm{DDIM}(\boldsymbol{x}_t, F^\theta(\boldsymbol{x}_t), t, r)
\end{aligned}
$$

(Flow Map Construction)

$$
\begin{aligned}
l_{\mathrm{ct}}(\boldsymbol{x}_t, r, s, t; \theta) &= \mathrm{LPIPS}(f^\theta(\boldsymbol{x}_t), \mathrm{sg}(f^\theta(\hat{\boldsymbol{x}}_s))), \\
&= \mathrm{LPIPS}(X^\theta_{t,r}(\boldsymbol{x}_t), \mathrm{sg}(\hat{X}_{s,r}(\hat{\boldsymbol{x}}_s))) \\
&= \mathrm{LPIPS}\Big(\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}^\theta_t(\boldsymbol{x}_t), t, r),\ \mathrm{sg}\big(\mathrm{DDIM}(\hat{\boldsymbol{x}}_s, \boldsymbol{v}^\theta_s(\hat{\boldsymbol{x}}_s), s, r)\big)\Big)
\end{aligned}
$$

(Learning Objective)

		CT
Diffusion basis†		
Flow path		Cosine
Network F^θ	Architecture	U-Net
	Output	$\boldsymbol{v}^\theta$
Flow map construction		
Time sampler		(note: $*$) $t = \frac{\pi}{2}\arctan([\sigma_{\max}^{1/\rho} + \frac{\tau}{K}(\sigma_{\min}^{1/\rho} - \sigma_{\max}^{1/\rho})]^\rho)$ $s = \frac{\pi}{2}\arctan([\sigma_{\max}^{1/\rho} + \frac{\tau+1}{K}(\sigma_{\min}^{1/\rho} - \sigma_{\max}^{1/\rho})]^\rho)$ $r = 0$, where $\tau \sim \mathcal{U}\{0, \dots, K-1\}$
Two-step target	1st-step ($\hat{\boldsymbol{x}}_s$)	$\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_{t\|0}, t, s)$
	2nd-step ($\hat{\boldsymbol{x}}_r$)	$\mathrm{DDIM}(\hat{\boldsymbol{x}}_s, \boldsymbol{v}^\theta_s, s, r)$
One-step prediction ($\boldsymbol{x}^\theta_r$)		$\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}^\theta_t, t, r)$
Training		
Loss metric d		LPIPS
sg EMA decay		✓

Examples: DTSC & CTSC

2. Shortcut Diffusion (DTSC)

$$\boldsymbol{x}_0 \sim p_0, \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(0,1)$$
$$\boldsymbol{x}_t = (1-t)\boldsymbol{x}_0 + t\boldsymbol{\varepsilon}$$
$$\boldsymbol{v}_{t|0} = -\boldsymbol{x}_0 + \boldsymbol{\varepsilon}$$

(Conditional Flow Path)

$$\hat{X}_{t,s}(\boldsymbol{x}_t) = \hat{\boldsymbol{x}}_s = \boldsymbol{x}_t - h\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t)$$
$$\hat{X}_{s,r}(\boldsymbol{x}_s) = \hat{\boldsymbol{x}}_r = \boldsymbol{x}_s - h\boldsymbol{u}^{\theta}_{s,r}(\hat{\boldsymbol{x}}_s)$$
$$X^{\theta}_{t,r}(\boldsymbol{x}_t) = \hat{\boldsymbol{x}}^{\theta}_t = \boldsymbol{x}_t - 2h\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)$$

(Flow Map Construction)

$$\begin{aligned} l_{\text{scd}}(\boldsymbol{x}_t, r, s, t; \theta) &= \frac{1}{4h^2} \cdot \left\| \boldsymbol{x}_t - 2h\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - \text{sg}\left(\boldsymbol{x}_t - h\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t) - h\boldsymbol{u}^{\theta}_{s,r}(\boldsymbol{x}_t - h\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t))\right) \right\|_2^2 \\ &= \frac{1}{4h^2} \cdot \left\| \boldsymbol{x}_t - 2h\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - \text{sg}\left(\boldsymbol{x}_t - h\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t) - h\boldsymbol{u}^{\theta}_{s,r}(\hat{\boldsymbol{x}}_s)\right) \right\|_2^2 \\ &= \left\| \boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - \frac{1}{2}\text{sg}\left(\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t) + \boldsymbol{u}^{\theta}_{s,r}(\hat{\boldsymbol{x}}_s)\right) \right\|_2^2, \end{aligned}$$

(Learning Objective)

		SCD
Diffusion basis†		
Flow path		Linear
Network F^{θ}	Architecture	DiT
	Output	$\boldsymbol{u}^{\theta}$
Flow map construction		
Time sampler		$t = \tau$ $s = \tau - h$ $r = \tau - 2h$ and with p_{teq}, $r = s = t$, where $\tau, h \sim$ Uniform $\log_2(\tau, h)$
Two-step target	1st-step ($\hat{\boldsymbol{x}}_s$)	$\boldsymbol{x}_t - h\boldsymbol{u}^{\theta}_{t,s}(\boldsymbol{x}_t)$
	2nd-step ($\hat{\boldsymbol{x}}_r$)	$\hat{\boldsymbol{x}}_s - h\boldsymbol{u}^{\theta}_{s,r}(\hat{\boldsymbol{x}}_s)$
One-step prediction ($\boldsymbol{x}^{\theta}_r$)		$\boldsymbol{x}_t - 2h\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)$
Training		
Loss metric d		Squared l_2-distance
sg EMA decay		✗

Examples: DTSC & CTSC

3. MeanFlow (CTSC)

Consistency property

$$\begin{aligned}&\left(\boldsymbol{x}_t + (r-t)\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)\right) - \left(\boldsymbol{x}_t + (s-t)\boldsymbol{v}_t + (r-s)\boldsymbol{u}^{\theta}_{s,r}(\boldsymbol{x}_s)\right)\\ =&(r-t)\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - (s-t)\boldsymbol{v}_t - (r-s)\boldsymbol{u}^{\theta}_{s,r}(\boldsymbol{x}_s)\,.\end{aligned}$$

$s \to t$ and normalized by dt

$$\begin{aligned}&((r-t)\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) + dt\cdot\boldsymbol{v}_t - (r-t+dt)\boldsymbol{u}^{\theta}_{t-dt,r}(\boldsymbol{x}_{t-dt}))/dt\\ =&dt\cdot\left(\boldsymbol{v}_t + \frac{d\left[(r-t)\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)\right]}{dt}\right)/dt\\ =&\boldsymbol{v}_t - \boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - (t-r)\frac{d}{dt}\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t).\end{aligned}$$

Replace v_t with $v_{t|0}$

$$l(\boldsymbol{x}_t, r, t-dt, t; \theta) = w\cdot\left\|\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t) - \mathrm{sg}\left(\boldsymbol{v}_{t|0} + (r-t)\frac{d\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)}{dt}\right)\right\|^2$$

(Learning Objective)

		MeanFlow
Diffusion basis†		
Flow path		Linear
Network F^θ	Architecture	DiT
	Output	$\boldsymbol{u}^\theta$
Flow map construction		
Time sampler		$r, t = \{\mathrm{sigmoid}(\tau_1), \mathrm{sigmoid}(\tau_2)\}$ s.t. $r \le t$, $s = t - dt$, and with p_{teq}, $r = s = t$, where $\tau_1, \tau_2 \sim \mathcal{N}(P_{\text{mean}}, P^2_{\text{std}})$
		(note: ‡)
Two-step target	1st-step ($\hat{\boldsymbol{x}}_s$)	$\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_{t\|0}, t, s)$
	2nd-step ($\hat{\boldsymbol{x}}_r$)	$\hat{\boldsymbol{x}}_s + (r-s)\boldsymbol{u}^{\theta}_{s,r}(\boldsymbol{x}_s)$
One-step prediction ($\boldsymbol{x}^\theta_r$)		$\boldsymbol{x}_t + (r-t)\boldsymbol{u}^{\theta}_{t,r}(\boldsymbol{x}_t)$
Training		
Loss metric d		Squared l_2-distance
sg EMA decay		✗

Examples: DTSC & CTSC

4. sConsistency Training Model (CTSC)

Gradient approximation

$$\nabla_\theta l(\boldsymbol{x}_t, r, s, t; \theta) \approx \nabla_\theta \|\boldsymbol{v}_t^\theta(\boldsymbol{x}_t) - \mathrm{sg}(\boldsymbol{v}_t^\theta(\boldsymbol{x}_t) + w(t)\tfrac{d}{dt} X_{t,r}^\theta(\boldsymbol{x}_t))\|_2^2.$$

Replace with DDIM solver

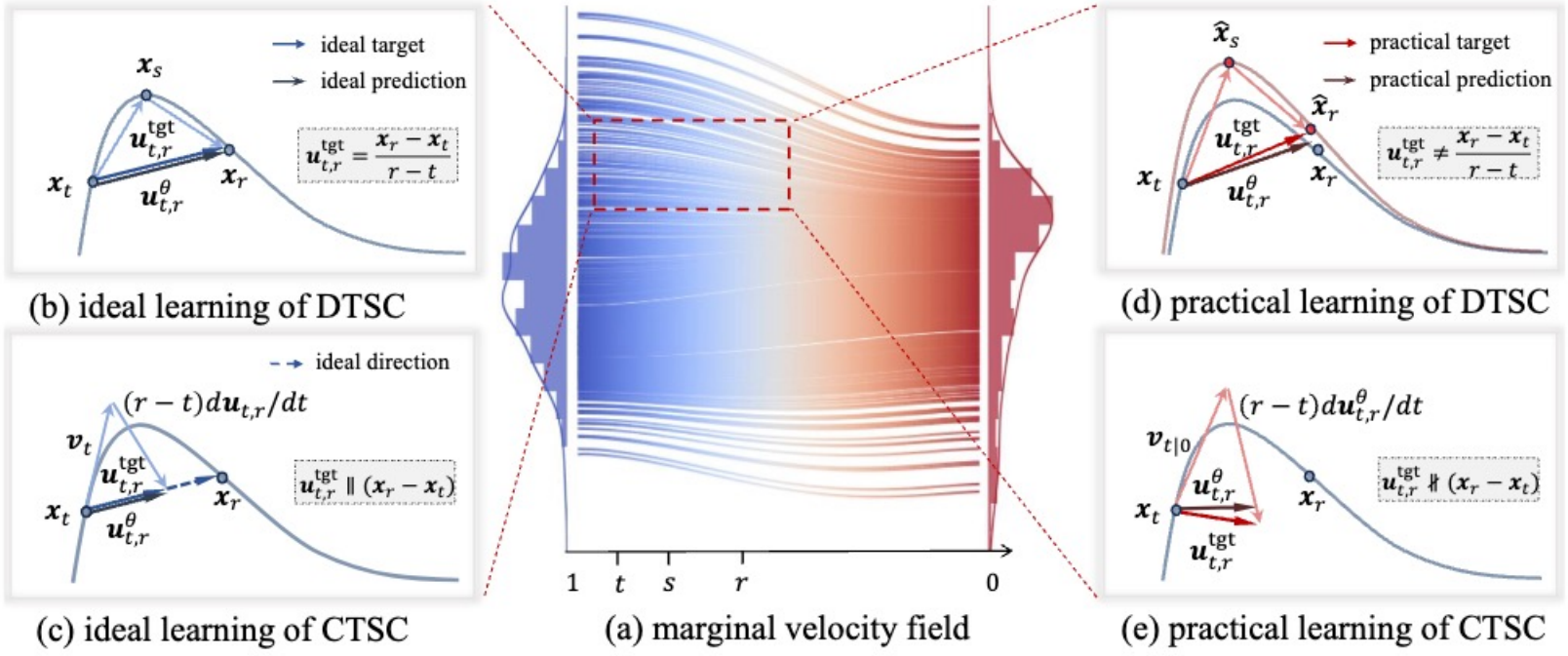
$$l_{\mathrm{sct}}(\boldsymbol{x}_t, r, t-dt, t; \theta) = \left\| \boldsymbol{v}_t^\theta(\boldsymbol{x}_t) - \mathrm{sg}\left(\boldsymbol{v}_t^\theta(\boldsymbol{x}_t) + w(t) \frac{d\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_t^\theta(\boldsymbol{x}_t), t, r)}{dt} \right) \right\|_2^2,$$

(Learning Objective)

sCT with linear paths is of the same form as MeanFlow

		sCT
Diffusion basis†		
Flow path		Cosine
Network F^θ	Architecture	U-Net
	Output	$\boldsymbol{v}^\theta$
Flow map construction		
Time sampler		$t = \frac{2}{\pi}\arctan(\exp(\tau))$ $s = t - dt$ $r = 0$, where $\tau \sim \mathcal{N}(P_{\mathrm{mean}}, P_{\mathrm{std}}^2)$
		(note: ‡)
Two-step target	1st-step ($\hat{\boldsymbol{x}}_s$)	$\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_{t\|0}, t, s)$
	2nd-step ($\hat{\boldsymbol{x}}_r$)	$\mathrm{DDIM}(\hat{\boldsymbol{x}}_s, \boldsymbol{v}_s^\theta, s, r)$
One-step prediction ($\boldsymbol{x}_r^\theta$)		$\mathrm{DDIM}(\boldsymbol{x}_t, \boldsymbol{v}_t^\theta, t, r)$
Training		
Loss metric d		Squared l_2-distance
sg EMA decay		✗

Discussion



(b) ideal learning of DTSC

(d) practical learning of DTSC

(c) ideal learning of CTSC

(a) marginal velocity field

(e) practical learning of CTSC

Why sharing a common design frame?

What challenges in constructing flow map targets?

Why distillation from pretrained velocity fields performs better?

Error bound of DTSC&CTSC

Under the assumption:

(i) one-sided Lipschitz condition of marginal velocity

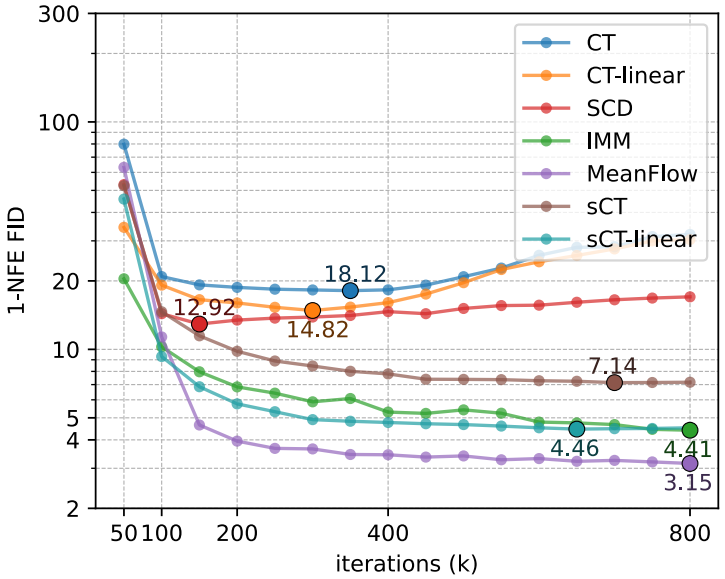
(ii) twice continuous differentiability with bounded second derivatives of $X^{\theta}_{\tau_1,\tau_2}$

$$W_2^2(p_0, p_0^\theta) \le C_1 \mathcal{L}_{dtsc}(\theta) + C_2(t-s);$$

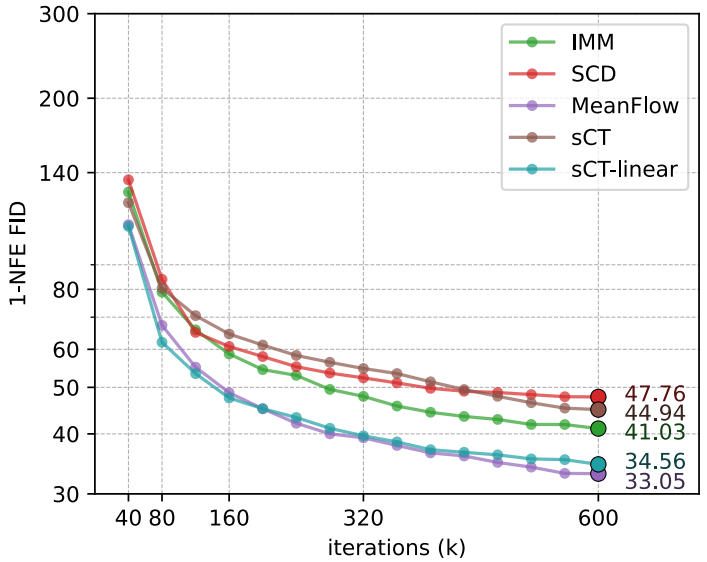
$$W_2^2(p_0, p_0^\theta) \le C_3 \mathcal{L}_{ctsc}(\theta),$$

$W_2(\cdot,\cdot)$ is the Wasserstein-2 distance.

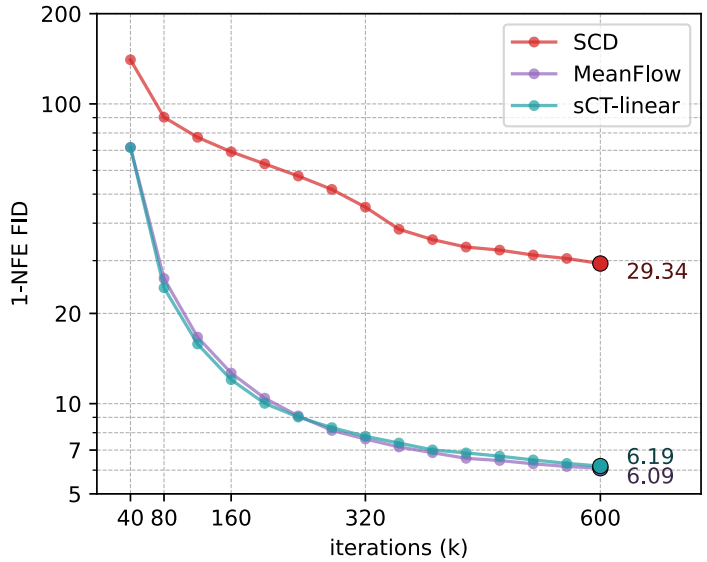
Elucidating the Design Space



Uncond. CIFAR



Cond. ImageNet



CFG. ImageNet

Following a linear or cosine path?

Shortcutting flow paths discretely or continuously?

Fixing the terminal time or not?

1. Linear paths in the setting of shortcut models are optimal under Fisher information metrics in the setting of shortcut models

2. The variance of conditional velocity $\sigma^2_{v_{t|0}} := \mathrm{Var}(\boldsymbol{v}_t(\boldsymbol{x}_t|\boldsymbol{x}_0))$ is the key to training stability and inference fidelity

Improvements to Training

Table 2: Evaluation of training improvements under one-step generation with SiT-B/2 as F^{θ}.

	Training configuration	FID50k
	MeanFlow under CFG. (Baseline)	6.09
+A1	Plug-in velocity ($p_{\text{plugin}} = 1.0$)	6.01
+A2	Plug-in velocity ($p_{\text{plugin}} = 0.5$)	5.98
+B1	Plug-in velocity ($p_{\text{plugin}} = 1.0$) & class-consistent batching	6.08
+B2	Plug-in velocity ($p_{\text{plugin}} = 0.5$) & class-consistent batching	5.96
+C	Gradual time sampler	5.99
+D	sCM training techniques	5.95
	ESC (Baseline + B2 + C + D)	**5.77**

Algorithm 1 Calculation of Plug-in Velocity

```
# x: training batch (B,D)
# t: sampled time
e = randn_like(x)
xt = (1- t) * x + t * e
x_ex, xt_ex = x[:,None,:], xt[None,:,:]
eps = (xt_ex - (1- t) * x_ex) / t

logp_fn = Normal(0, 1).log_prob
logp = sum(logp_fn(eps), dim=2)
weight = softmax(logp, dim=0)

v_cnd = eps - x_ex
v_plugin = matmul(weight.T, v_cnd)
```

Use plug-in velocity instead

Ideal velocity

Assume the data distribution is the empirical distribution, as

$p_0(\boldsymbol{y}) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\boldsymbol{y}_i}(\boldsymbol{y})$,

The marginal velocity reads

$$\boldsymbol{v}_t^*(\boldsymbol{x}_t|\{\boldsymbol{y}^{(i)}\}_{i=1}^N) = \sum_i^N \frac{\mathcal{N}(\boldsymbol{x}_t; \alpha_t \boldsymbol{y}^{(i)}, \sigma_t^2 \mathbf{I})}{\sum_j^N \mathcal{N}(\boldsymbol{x}_t; \alpha_t \boldsymbol{y}^{(j)}, \sigma_t^2 \mathbf{I})} \left(\dot{\alpha}_t \boldsymbol{y}^{(i)} + \frac{\dot{\sigma}_t}{\sigma_t}(\boldsymbol{x}_t - \alpha_t \boldsymbol{y}^{(i)})\right),$$

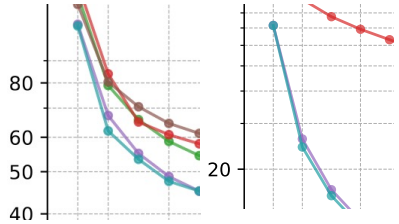
Plug-in velocity instead of conditional one:

- Reducing the level of variance $\sigma_{v_{t|0}}$ to $O(1/B)$, at the minor cost of increased bias
- Employing class-consistent mini-batching in guidance training to avoiding dilution the class specific signal

Gradual time sampler from sCT to MeanFlow:

- Fast convergence to local minima around the global one

Adoption of training techniques from sCT



Scaling up

Table 3: Evaluation on ImageNet-256×256. Values with underline denote the best except shortcut models, values in **bold** is the best shortcut diffusion model under one-step generation.

Family	Method	Param.	NFE	FID50k
GAN	BigGAN (Brock et al., 2019)	112M	1	6.95
	GigaGAN (Kang et al., 2023)	569M	1	3.45
	StyleGAN-XL (Karras et al., 2019)	166M	1	2.30
AR/Mask	AR w/ VQGAN (Esser et al., 2021)	227M	1024	26.52
	MaskGIT (Chang et al., 2022)	227M	8	6.18
	VAR-d30 (Tian et al., 2024)	2B	10×2	1.92
	MAR-H (Li et al., 2024)	943M	256×2	1.55
Diff/ Flow	ADM (Karras et al., 2024)	554M	250×2	10.94
	LDM-4-G (Rombach et al., 2021)	400M	250×2	3.60
	SimDiff (Hoogeboom et al., 2023)	2B	512×2	2.77
	DiT-XL/2 (Peebles & Xie, 2022)	675M	250×2	2.27
	SiT-XL/2 (Ma et al., 2024)	675M	250×2	2.06
	SiT-XL/2+REPA (Yu et al., 2025)	675M	250×2	1.42
Shortcut	iCT (Song & Dhariwal, 2023)	675M	1	34.24
	SCD (Frans et al., 2025)	675M	1	10.60
	IMM (Zhou et al., 2025)	675M	1×2	7.77
	MeanFlow (Geng et al., 2025a)	676M	1	3.43
			2	2.93
	ESC (w/o-class-consist.)	676M	1	2.92
	ESC (w/-class-consist.)	676M	1	2.85
	ESC+ (w/-class-consist.)	676M	1	**2.53**

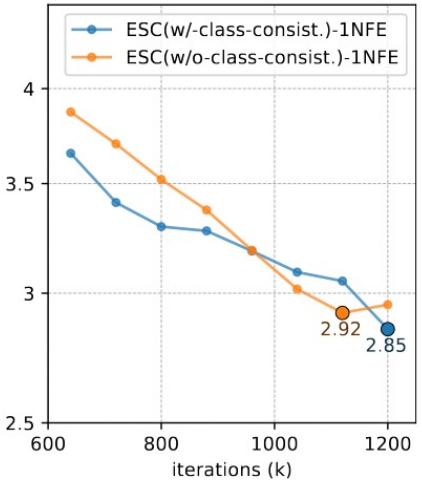
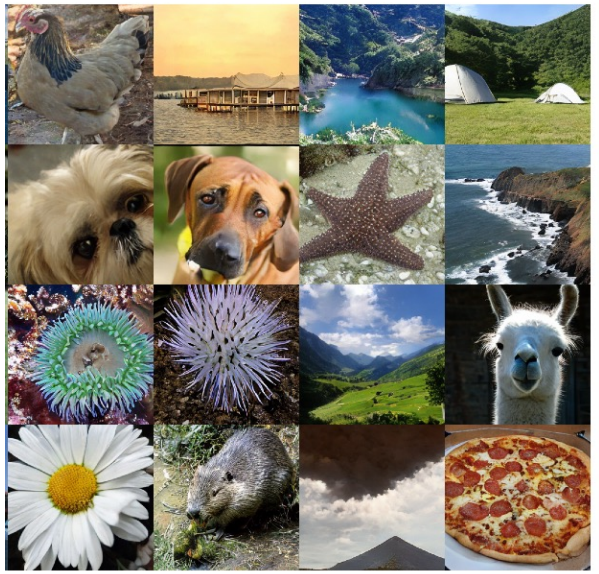


Figure 3: Convergence of FID50k.

Table 4: Uncond. CIFAR-10.

method	NFE	FID
iCT	1	**2.83**
ECT	1	3.60
sCT	1	2.97
IMM	1	3.20
MeanFlow	1	2.92
ESC	1	**2.83**

- **State-of-the-art Performance**
- **Minimal time cost**
- **Faster convergence**



ESC with SiT-XL/2 (2.85 FID)

Thanks!

Center of AI for Research and Innovation, School of Engineering, Westlake University;

AI for Scientific Simulation and Discovery Lab, School of Engineering, Westlake University;

Academy of Mathematics and Systems Science (AMSS) of the Chinese Academy of Sciences;

DP Technology

Paper:

Haitao Lin, Peiyan Hu*, Minsi Ren, Zhifeng Gao, Zhi-Ming Ma, Guolin Ke, Tailin Wu, Stan Z. Li* ,ON THE DESIGN OF ONE-STEP DIFFUSION VIA SHORTCUTTING FLOW PATHS, url: https://openreview.net/pdf?id=k6q8rRYVQR

Code:

https://github.com/EDAPINENUT/ExplicitShortCut