Constant Acceleration Flow

Dogyun Park, Sojin Lee, Sihyeon Kim, Taehoon Lee, Youngjoon Hong[†], Hyunwoo J. Kim[†]

MLV Machine Learning & Vision Lab · KOREA UNIVERSITY · KAIST

Github · Paper

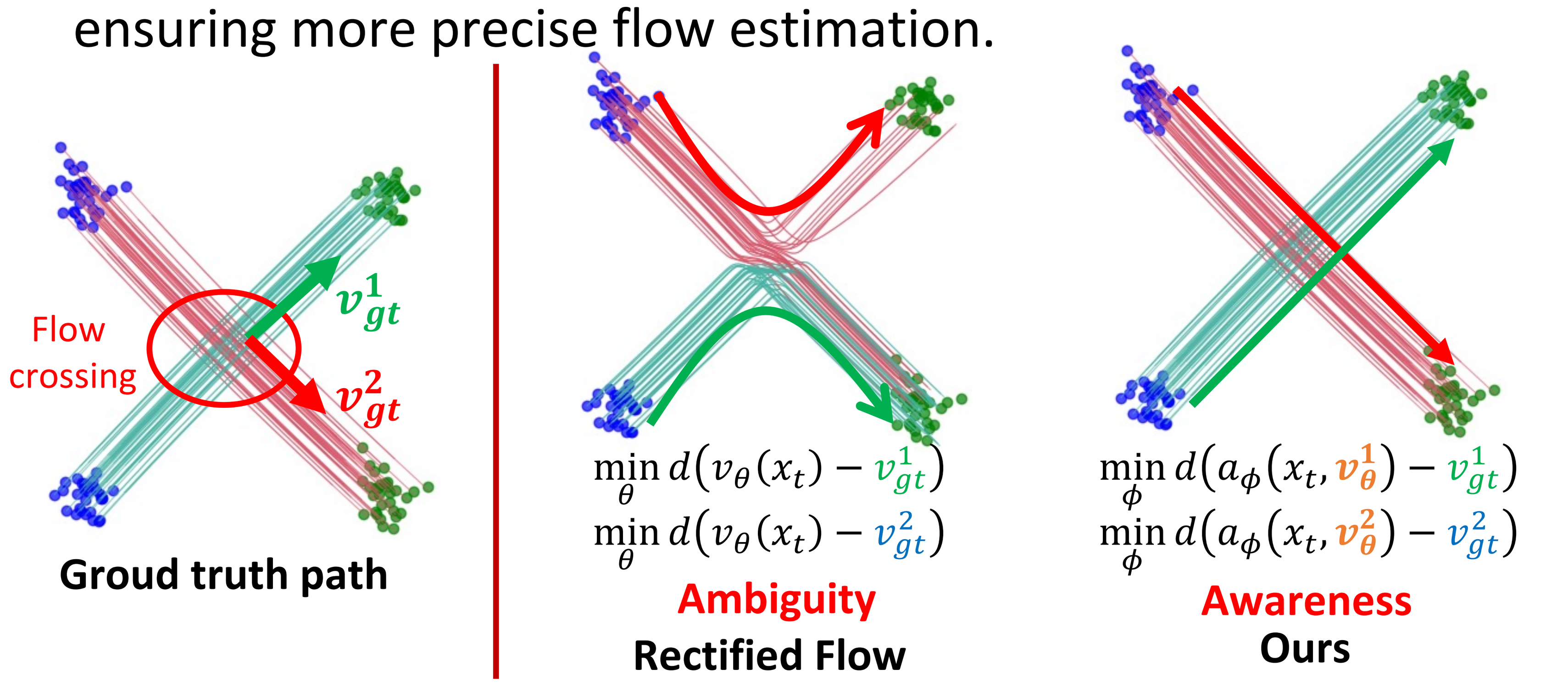
NEURAL INFORMATION PROCESSING SYSTEMS

Motivation

- **Flow-based approaches**, such as **rectified flow/reflow**, have demonstrated remarkable success in few-step generation.
- However, their performance remains limited in few-step scenarios, due to **two key challenges**:
  1) **Ambiguity**: **Flow crossing** introduce directional ambiguity, leading to **estimation inaccuracies**.
  2) **Expressivity**: Modeling flows between complex distributions with a single velocity **may limit expressivity** to capture intricate patterns.

Flow crossing

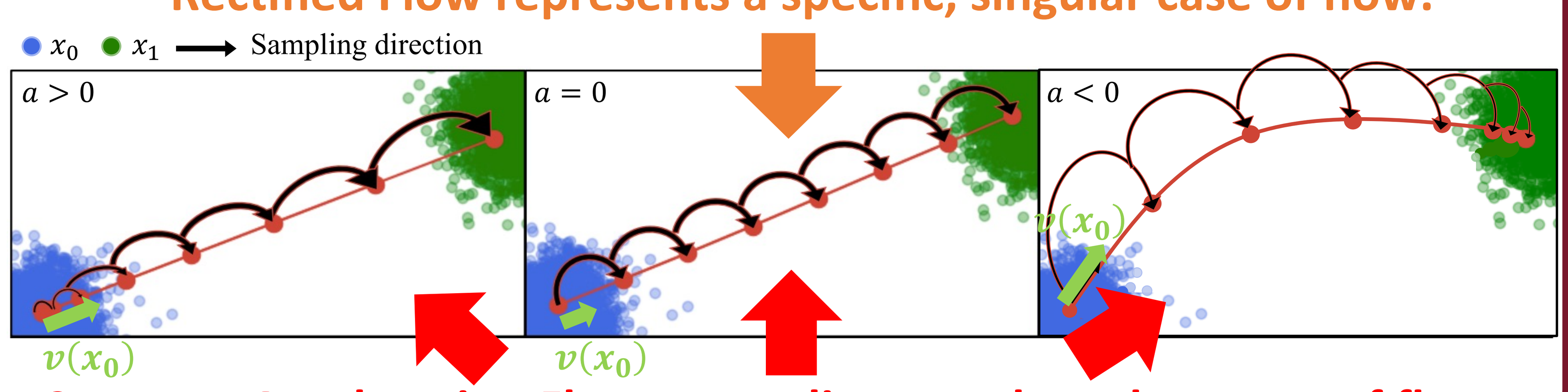
- **Flow crossing** $(x_t^1 = x_t^2)$ results in different ground truth targets at the same location, introducing **ambiguity** in learning.
- This ambiguity causes flows to **curve**, reducing accuracy in few-step sampling.
- Our **Initial Velocity Conditioning** mitigates this limitation, ensuring more precise flow estimation.



Constant Velocity vs. Constant Acceleration

- Rectified flow only represents **linear flow with constant speed**.
- Constant Acceleration Flow can represent **diverse flows** based on the **initial velocity $v_0(x_0)$ with closed-form solution.**

Rectified Flow represents a specific, singular case of flow.



Constant Acceleration Flow generalizes to a broader range of flows.

Main framework

✓ **Ordinary Differential Equation** of Constant Acceleration Flow

Eq(1). CAF ODE

$$dx_t = \underbrace{v(x_0)}_{\text{Initial Velocity}}dt + t \cdot \underbrace{a(x_t)}_{\text{Acceleration field}}dt$$

- By integrating both sides of Eq(1) w.r.t time and assuming a **constant acceleration** field $(\boldsymbol{a}(\boldsymbol{x}_{t_1}) = \boldsymbol{a}(\boldsymbol{x}_{t_2}), \forall \boldsymbol{t}_1, \boldsymbol{t}_2 \in [\mathbf{0}, \mathbf{1}])$, we derive the following **solution of ODE**:

$$x_t = x_0 + v(x_0)t + \frac{1}{2}a(x_t)t^2 \xrightarrow{t=1} x_1 = x_0 + v(x_0) + \frac{1}{2}a(x_t)$$

Eq(2). Closed-form solution — *Single-step sampling!*

✓ **Stage 1. Initial Velocity Field v_θ**

- The **initial velocity** is defined as a **scaled displacement vector** between x_1 and x_0.
- θ is optimized to minimize a **distance metric d** between target and estimation.

$$\underbrace{v(x_0)}_{\text{Initial Velocity}} = h(x_1 - x_0) \Rightarrow \min_\theta \mathbb{E}\left[d(v(x_0), v_\theta(x_t))\right]$$

✓ **Stage 2. Acceleration Field a_ϕ**

- Using the learned initial velocity field v_θ, the corresponding **acceleration field** is derived directly from Eq(2).

$$\underbrace{a(x_t)}_{\text{Acceleration field}} = 2(x_1 - x_0) - 2v_\theta(x_0) \Rightarrow \min_\phi \mathbb{E}\left[d\left(a(x_t), a_\phi(x_t, \underbrace{v_\theta(x_t)}_{\text{Initial Velocity Conditioning (IVC)}})\right)\right]$$

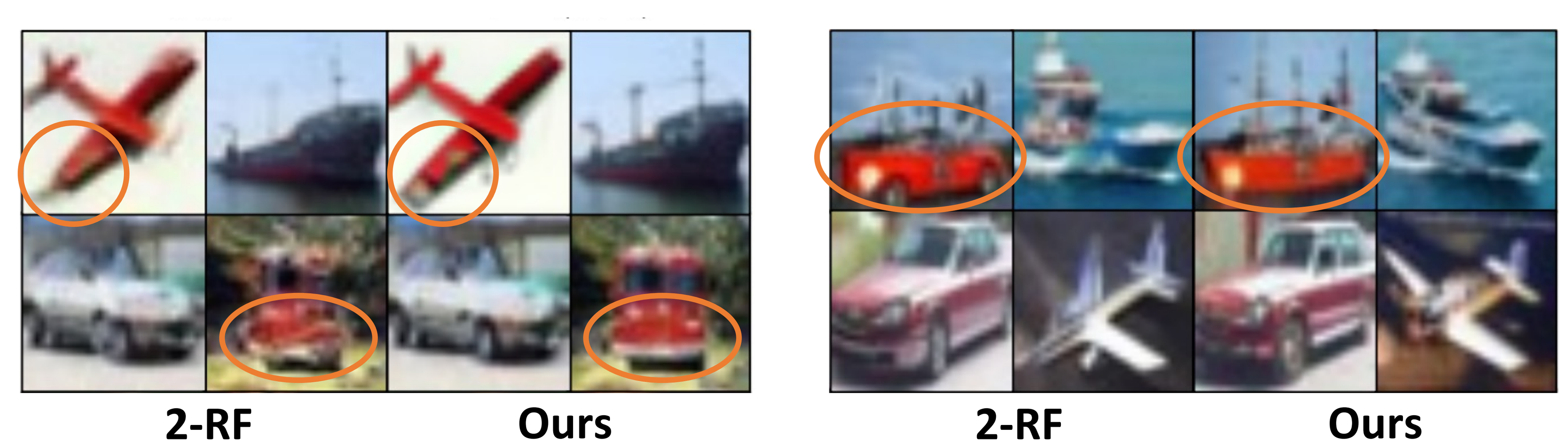
✓ **Initial Velocity Conditioning (IVC)**

- We introduce **conditioning the initial velocity** as an additional input to the acceleration model.
- This provides **directional information** to the model, effectively reducing ambiguity in flow estimation.

Qualitative results

✓ **Qualitative comparison** between 2-Rectified Flow and ours.

- Our model generates **more vivid and detailed images** than 2-RF.



Quantitative results

✓ CAF achieve **comparable or stronger performance** compared to SOTA models.

CIFAR10 32x32

Diffusion/Consistency Models	N	FID ↓	FID ↓
Diff-Instruct [9]	1	4.53	-
DMD [44]	1	3.77	-
DFNO [5]	1	3.78	-
TRACT [45]	1	3.78	-
KD [46]	1	9.36	-
CD [6]	2	2.93	-
	1	3.55	-
CTM [7]	2	1.87	1.63
	1	1.98	1.73
Rectified Flow Models			
2-Rectified Flow [10]	2	7.89	3.74
	1	11.81	6.88
2-Rectified Flow + Distill [10]	1	4.84	-
CAF (Ours)	1	4.81	2.68
CAF + GAN (Ours)*	1	**1.48**	**1.39**

ImageNet 64x64

Diffusion/Consistency Models	N	FID ↓	IS ↑	Recall↑
Diff-Instruct [9]	1	5.57	-	-
DMD [44]	1	2.62	-	-
TRACT [45]	1	7.43	-	-
DFNO [5]	1	7.83	-	0.61
PD [3]	1	15.39	-	0.62
CD [6]	2	4.70	-	0.64
	1	6.20	40.08	0.57
CTM [7]	2	1.73	64.29	0.57
	1	1.92	70.38	0.57
Rectified Flow Models				
CAF (Ours)	1	6.52	37.45	0.62
CAF + GAN (Ours)*	1	**1.69**	62.03	0.64

*Finetuning with adversarial loss using real data

Analysis

Ablation study

✓ We conduct an ablation study to analyze the impact of **three components** in few-step generation:

Table 5: Ablation study on CIFAR-10 ($N = 1$).

Config	Constant acceleration	v_0 condition	Reflow procedure	FID↓
A	✗	✗	✗	378
B	✗	✗	✔	6.88
C	✔(h=1.5)	✗	✔	3.82
D	✔(h=1.5)	✔	✔	**2.68**

- **A vs. B**: Effectiveness of **reflow**
- **B vs. C**: Expressiveness of **CAF**
- **C vs. D**: Effectiveness of **IVC**

Applications

✓ **Reconstruction using CAF Inversion**

- By our IVC, CAF achieves accurate reconstruction using only a **single-step inversion.**

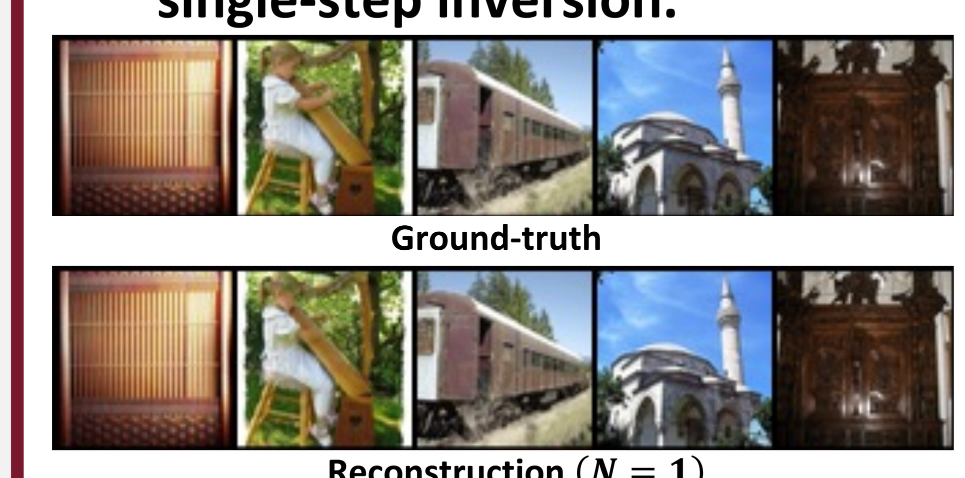


Table 6: Reconstruction error.

Model	N	PSNR ↑	LPIPS ↓
CM	-	N/A	N/A
CTM	-	N/A	N/A
EDM	4	13.85	0.447
2-RF	2	33.34	0.094
2-RF	1	29.33	0.204
CAF (Ours)	1	**46.68**	**0.007**
CAF (+GAN) (Ours)	1	40.84	0.028

✓ **Zero-shot Box Inpainting**