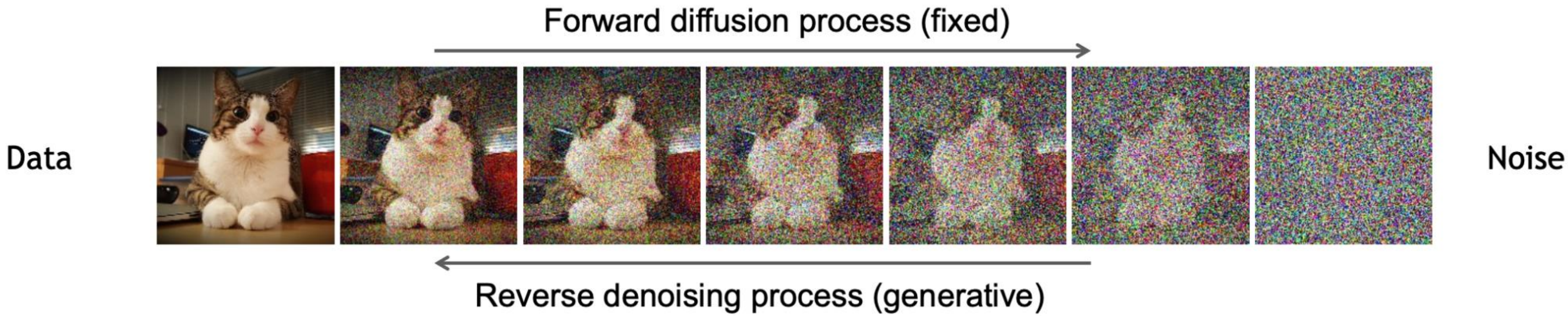
Improving the Training of Rectified Flows

Sangyun Lee[1], Zinan Lin[2], Giulia Fanti[1]

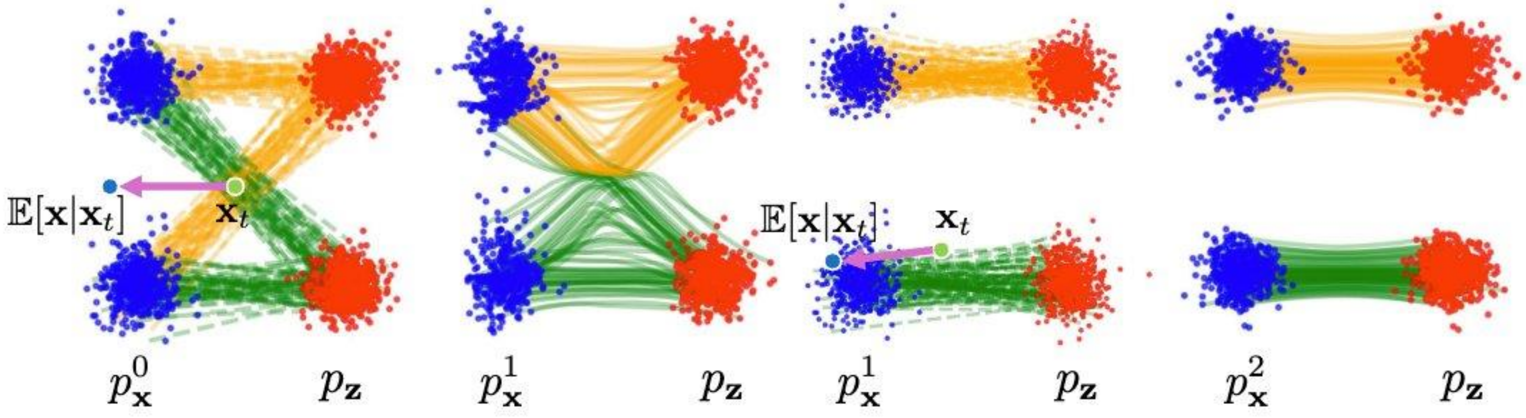
[1]Carnegie Mellon University, [2]Microsoft Research

Diffusion models are slow to sample from



https://cvpr2022-tutorial-diffusion-models.github.io/

Rectified flows speed up sampling by learning straight trajectories.



(a) Linear interpolation induced by p_{xz}^0

(b) Generative process of 1-rectified flow

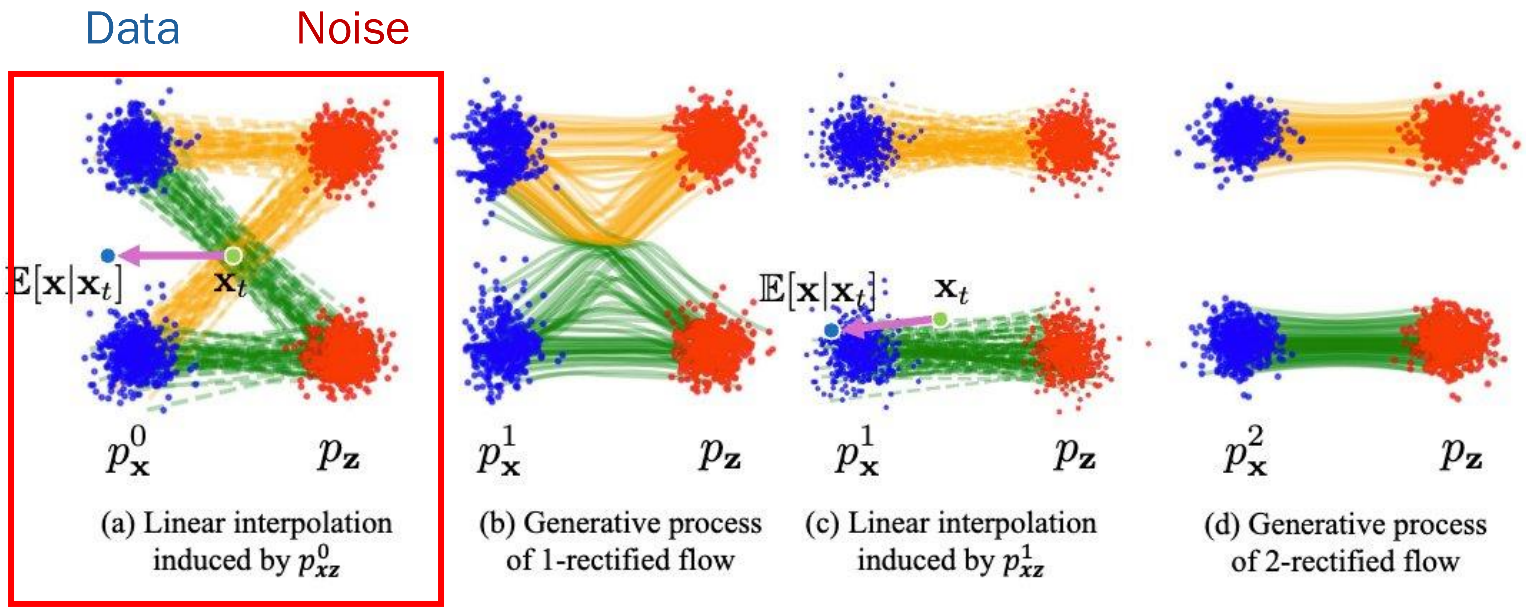
(c) Linear interpolation induced by p_{xz}^1

(d) Generative process of 2-rectified flow

Intro to Rectified Flows

Step 1: Independently sample noise-data pairs

Step 2: Linearly interpolate between noise-data pairs



(a) Linear interpolation induced by p_{xz}^0

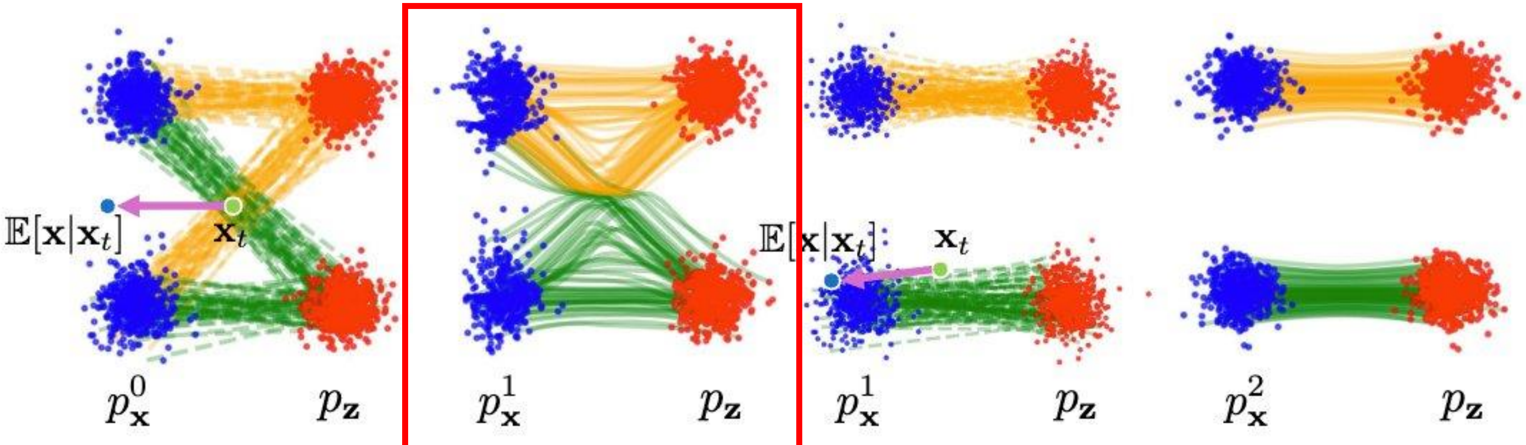
(b) Generative process of 1-rectified flow

(c) Linear interpolation induced by p_{xz}^1

(d) Generative process of 2-rectified flow

Intro to Rectified Flows

Step 3: Learn an ODE that with crossing-free trajectories



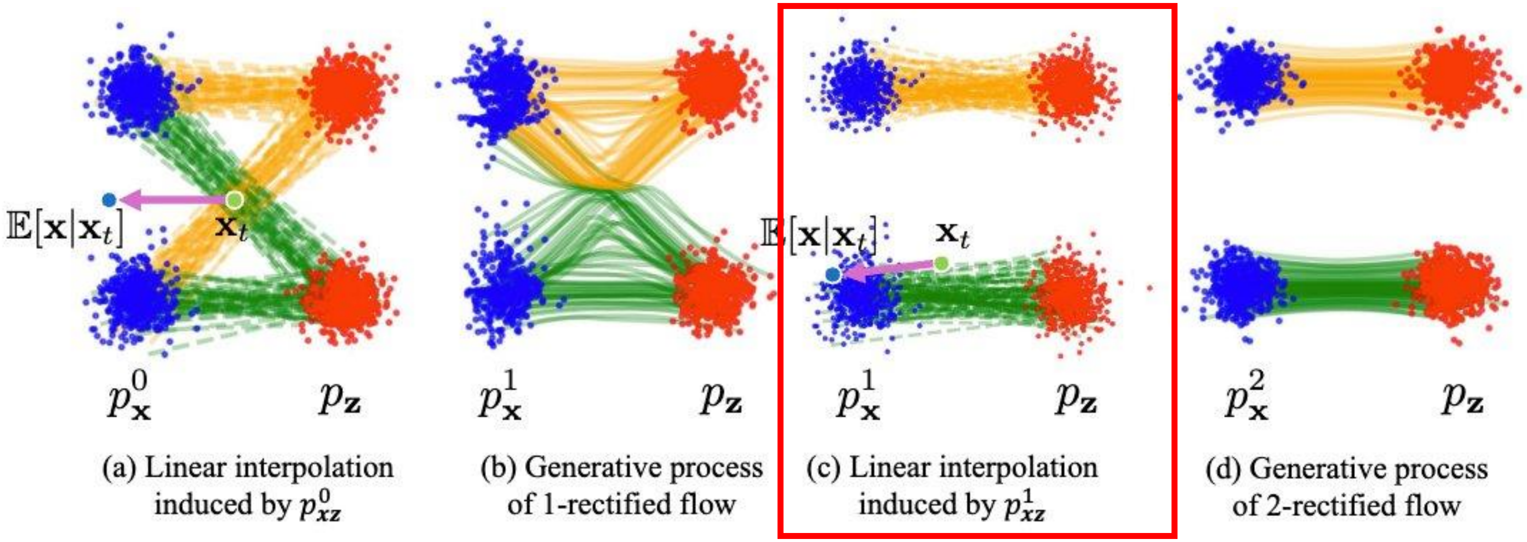
(a) Linear interpolation induced by p^0_{xz} (b) Generative process of 1-rectified flow (c) Linear interpolation induced by p^1_{xz} (d) Generative process of 2-rectified flow

Closed form ODE:

$$\frac{d\mathbf{z}_t}{dt} = \mathbf{v}_t(\mathbf{z}_t) := \frac{1}{t}(\mathbf{z}_t - \mathbb{E}[\mathbf{x}|\mathbf{x}_t = \mathbf{z}_t]).$$

Intro to Rectified Flows

Step 4: Generate noise-data pairs from new ODE

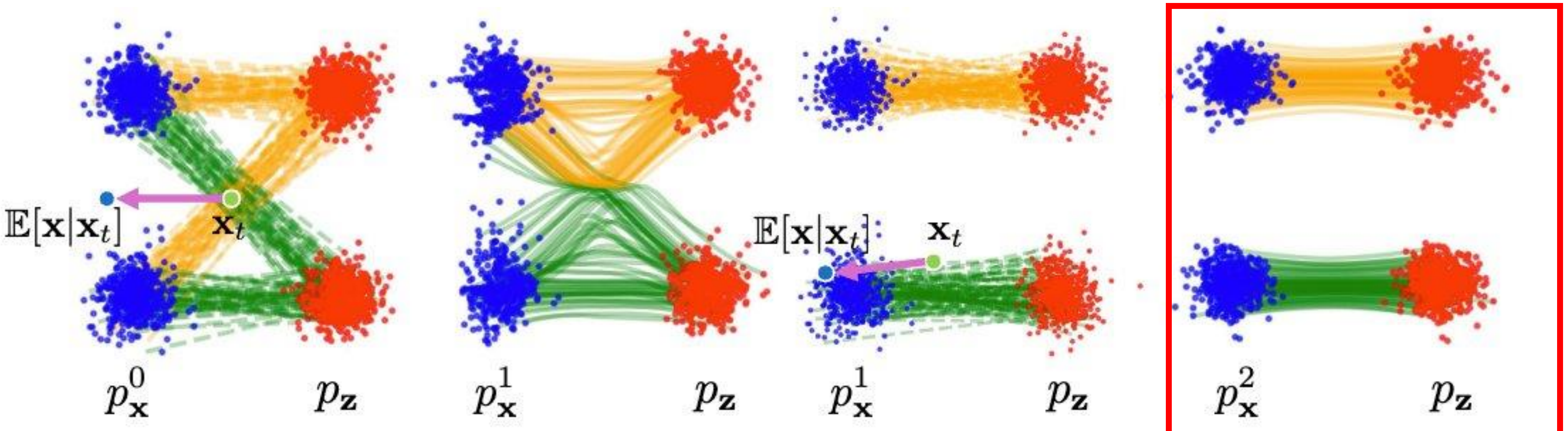


k-rectified flow = k rounds of repetition

Intro to Rectified Flows

Step 4: Generate noise-data pairs from new ODE

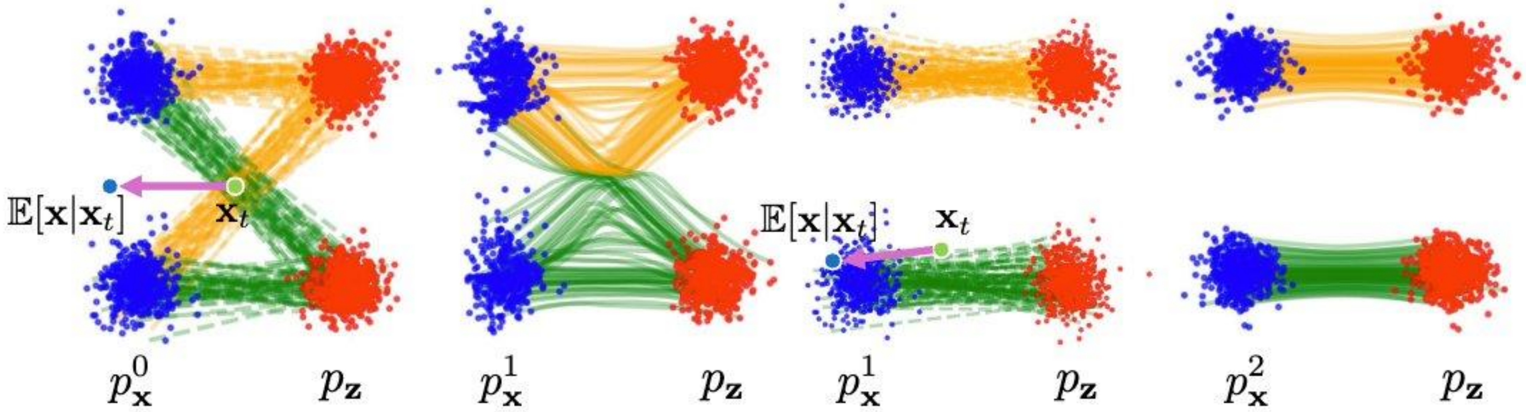
Step 5: Repeat



(a) Linear interpolation induced by p^0_{xz} (b) Generative process of 1-rectified flow (c) Linear interpolation induced by p^1_{xz} (d) Generative process of 2-rectified flow

k-rectified flow = k rounds of repetition

How many rounds are enough?



(a) Linear interpolation induced by p_{xz}^0 (b) Generative process of 1-rectified flow (c) Linear interpolation induced by p_{xz}^1 (d) Generative process of 2-rectified flow

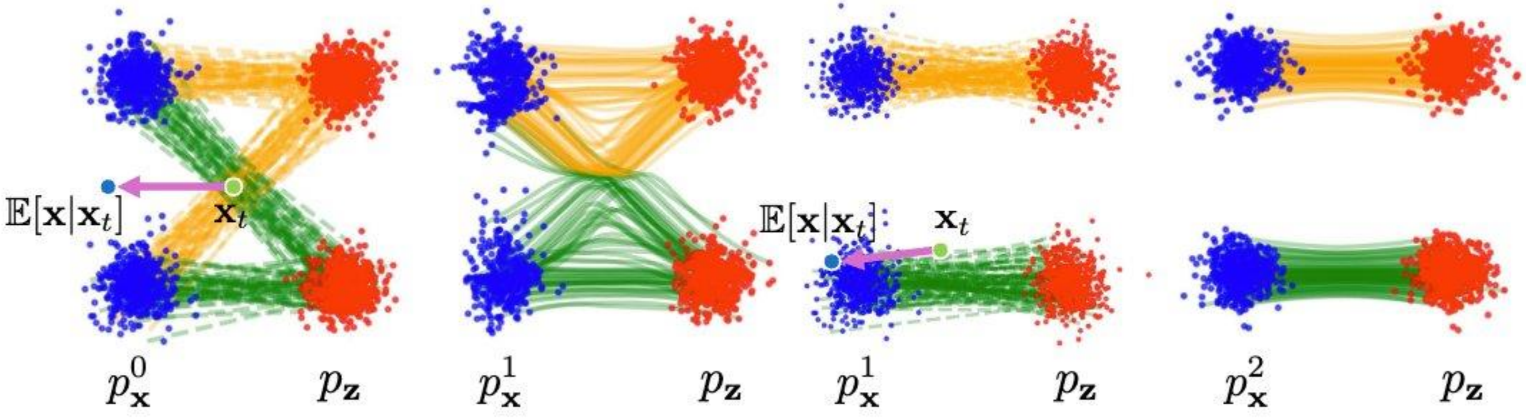
Theoretically
$k \to \infty$ gives straight trajectories

Conventionally
At least 3 rounds

Our claim:

Two rounds of Reflow are enough!

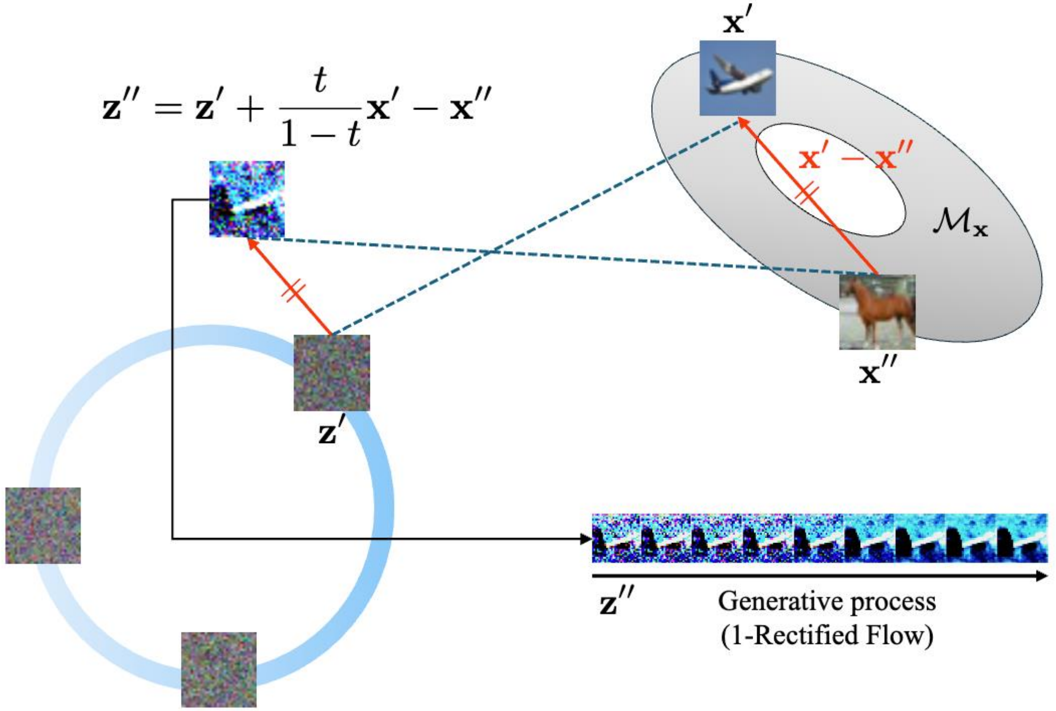
When do we get the straight ODE?



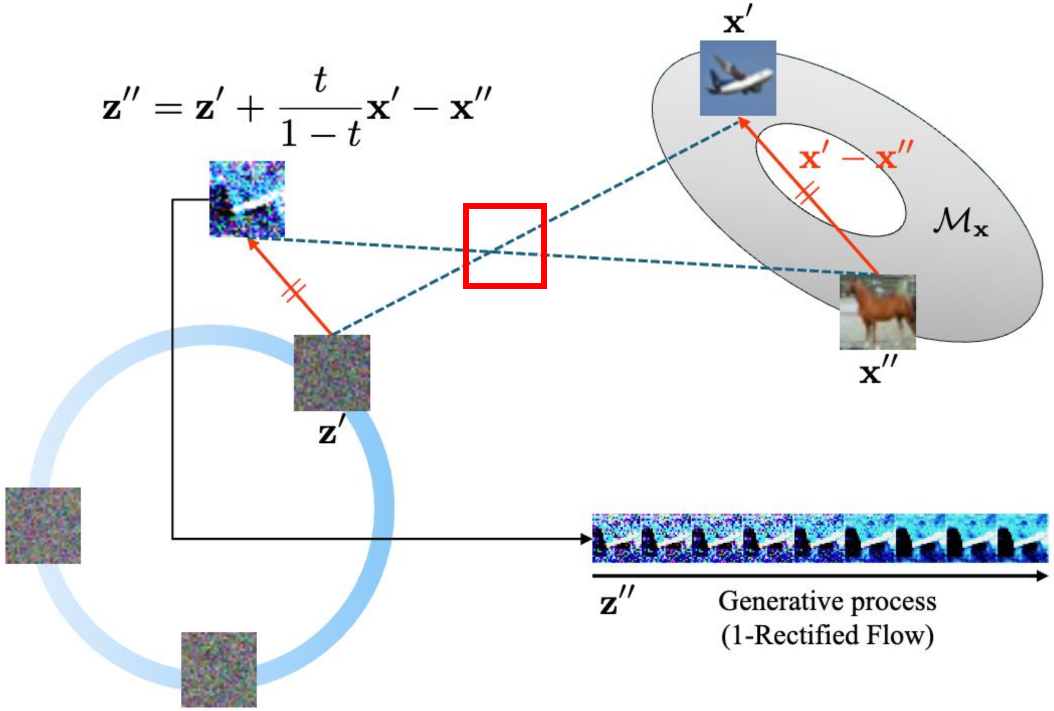
(a) Linear interpolation induced by p_{xz}^0 (b) Generative process of 1-rectified flow (c) Linear interpolation induced by p_{xz}^1 (d) Generative process of 2-rectified flow

k-rectified flow ODE is straight iff the linear interpolation paths of (k-1)-rectified flow do not intersect.

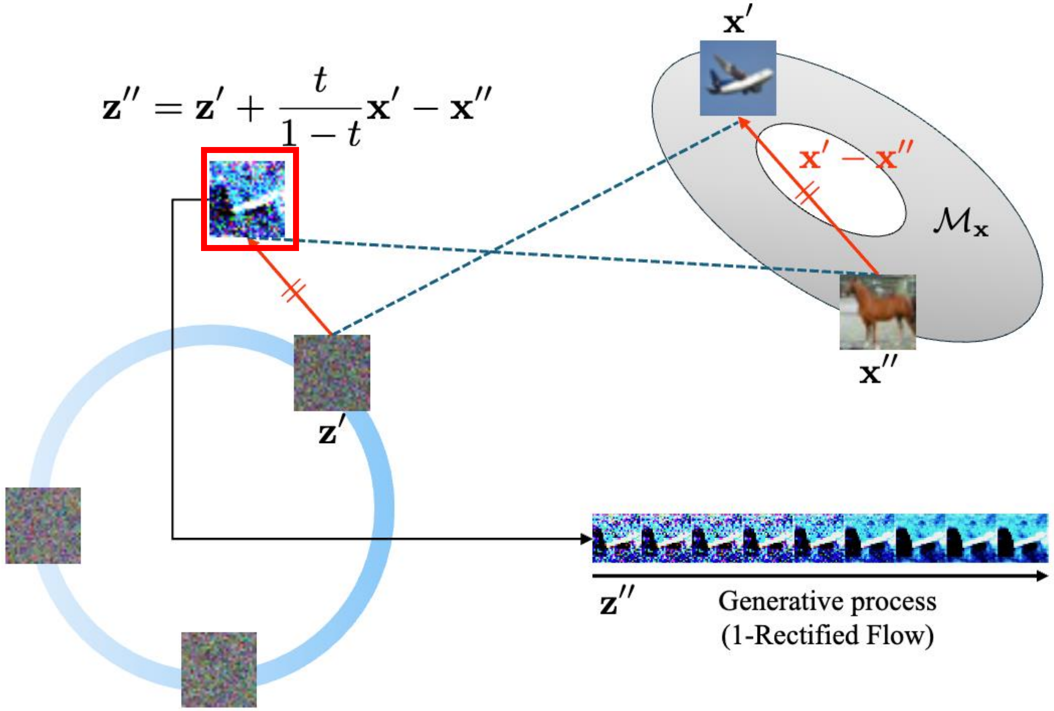
2-RF training is intersection-free.



2-RF training is intersection-free.

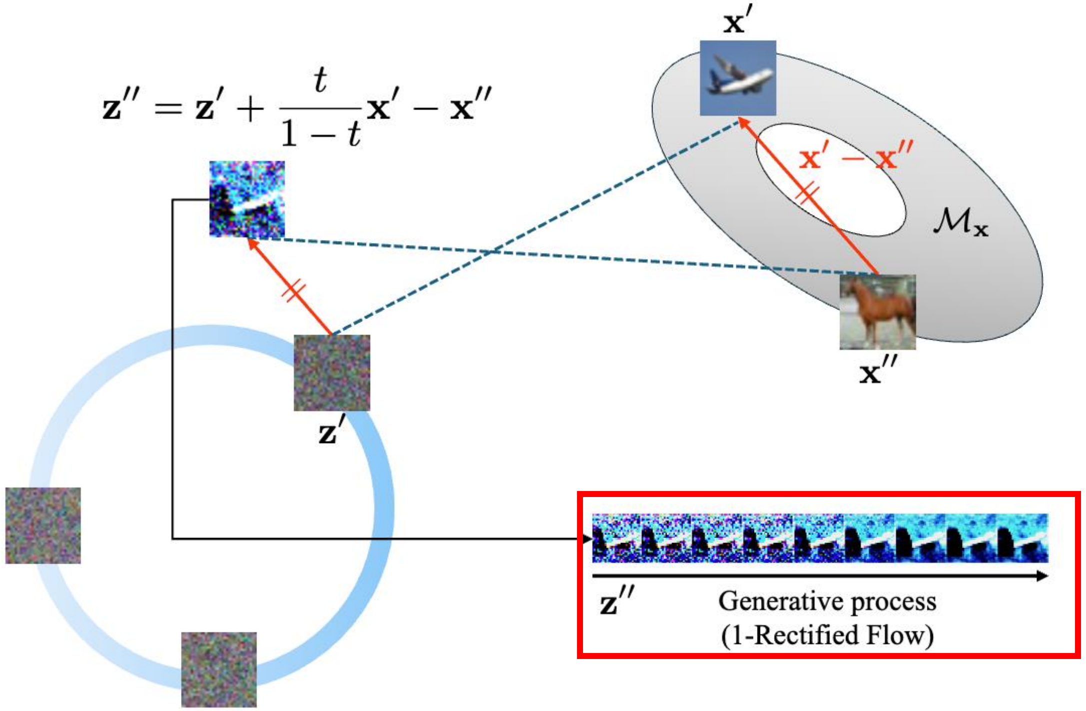


2-RF training is intersection-free.



Not a realistic Gaussian noise!

2-RF training is intersection-free.

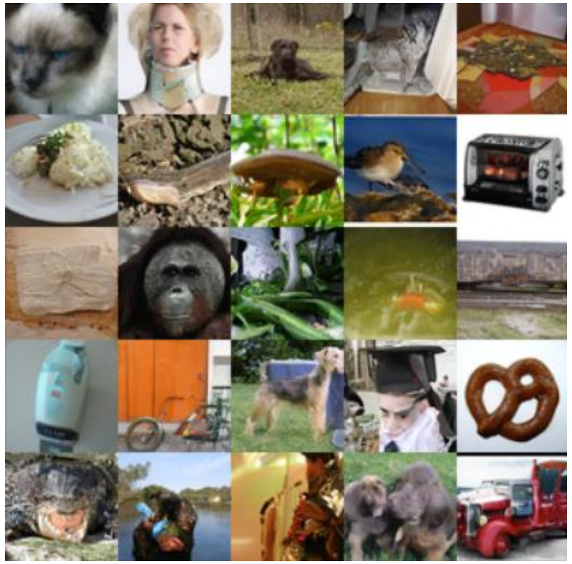


Cannot be mapped to a realistic horse image.

2-rectified flow++ is competitive with SOTA

FFHQ 64x64

AFHQ 64x64



ImageNet 64x64

One-step generated samples from our 2-rectified flow++.

In summary,

- The optimal 2-rectified flow is nearly straight.
  - This motivates new training techniques
- **Our improved techniques make 2-rectified flow competitive to SOTA methods.**
- It still retains and enhances the useful features of the neural ODE, such as multi-step iterative refinement and inversion from data to noise.

github.com/sangyun884/rfpp