Flow-Matching for Generative Modelling*
reproduction

Alexander Sharshavin
Artem Alexeev
Kseniia Petrushina
Daniil Shlenskii

Skoltech

2024.10.25

* https://arxiv.org/pdf/2210.02747

The essence of the work

If we have a flow $v(x_t, t)$, which preserves boundaries $x_0 \sim p_{prior}, x_1 \sim p_{data}$, then we can sample from the p_{data} by solving the following ODE:

$$\frac{d}{dt}x_t = v(x_t, t), \quad x_0 \sim p_{prior}$$

This work proposes a way to learn such a flow in a simulation-free manner (i.e. we do not have to approximate full ODE trajectory at training step).

The idea behind

Naïve way to learn the flow is using regression: $L_{FM}(\theta) = \mathbb{E}_{t,p_t(x)} \| v_\theta(x_t, t) - u(x_t, t) \|$

However, it is intractable!

But $u(x_t, t)$ can be expressed via conditional flow $u(x_t, t \,|\, x_1)$:

$$u(x_t, t) = \int u(x_t, t \,|\, x_1) \frac{p_t(x_t \,|\, x_1) q(x_1)}{p_t(x_t)} dx_1 = \int u(x_t, t \,|\, x_1) p_t(x_1 \,|\, x_t) dx_1 = \mathbb{E}_{x_1} \left\{ u(x_t, t \,|\, x_1) \,|\, x_t \right\}$$

Using properties of conditional expectation we get the following tractable Condition Flow Matching loss:

$$L_{CFM}(\theta) = \mathbb{E}_{t,q(x_1),p_t(x_t|x_1)} \| v_\theta(x_t, t) - u(x_t, t \,|\, x_1) \|^2$$

Our main interest is better generation

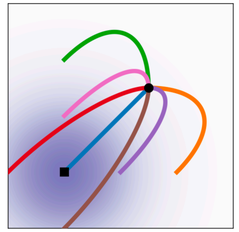
Current SotA generative models (score-based models)
struggle from time-consuming sampling process.

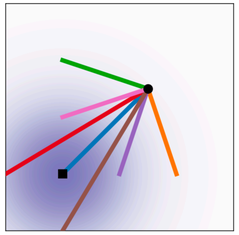
One of the reasons for that is high curvature of the backward trajectories.

Intuitively, if we would have perfectly straight trajectories,
we could solve ODE with one Euler step.

Authors set conditional flow in a way that ODE are straight:

$$x_t = tx_1 + (1-t)\varepsilon, \quad \epsilon \sim \mathcal{N}(0,I)$$





Main result

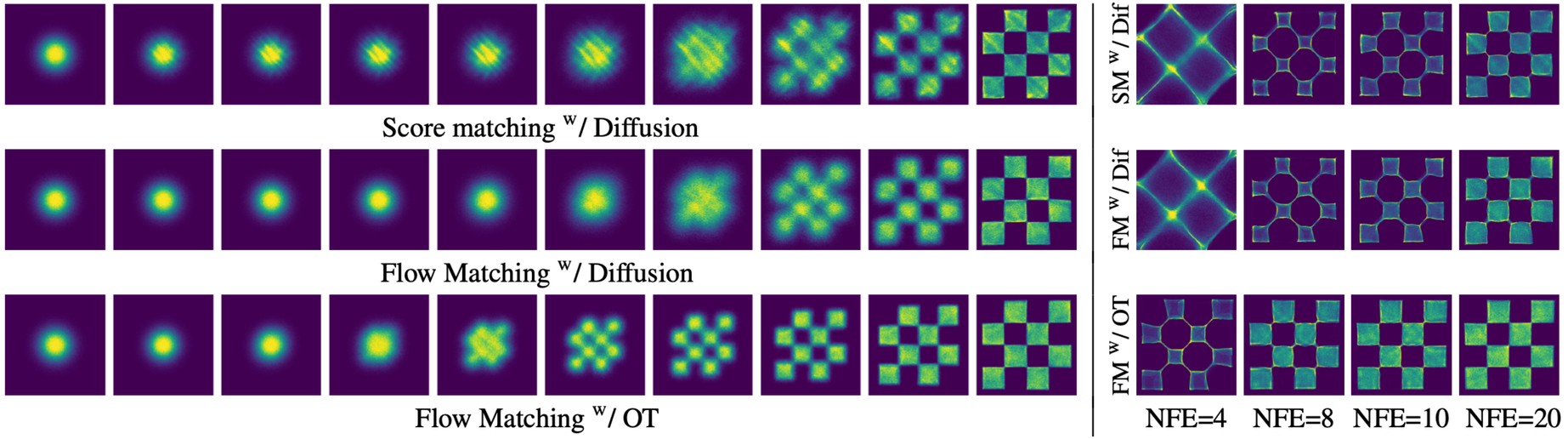
The main result of the work is that forced straighter trajectories induce faster and more quality sampling.

Checkerboard (our)

FM-OT

FM-Diffusion

Checkerboard (paper)



Checkerboard (our)

FM-OT

FM-Diffusion

Conclusion

1. FM-OT showed faster and more stable training

2. FM-OT enables faster sampling

3. However it is not really clear if straight trajectories are beneficial for sample quality in case of infinite resources. Actually there is work showing it is not*

* https://arxiv.org/abs/2404.12940

Some images

* https://arxiv.org/abs/2404.12940

Responsibilities

Alexander Sharashvin
- data collection
- method evaluation

Artem Alekseev
- training pipeline
- architectures

Kseniia Petrushina
- Conditional Flow Matching class
- inference

Daniil Shlenskii
- team managing
- slides + presentation

* https://arxiv.org/abs/2404.12940