香 港 大 學
THE UNIVERSITY OF HONG KONG

Flow Matching

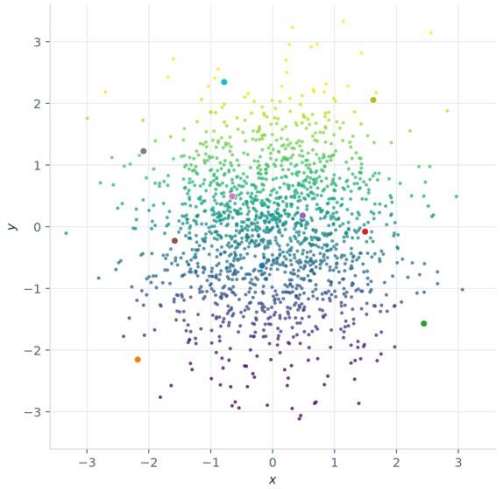
HKU ELEC8505 Tutorial Presentation

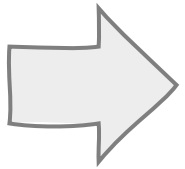
Yifeng Xu

2026.11.18

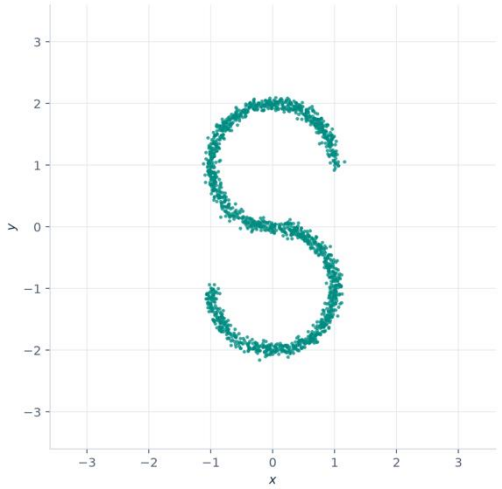
Generative Modeling

prior distribution



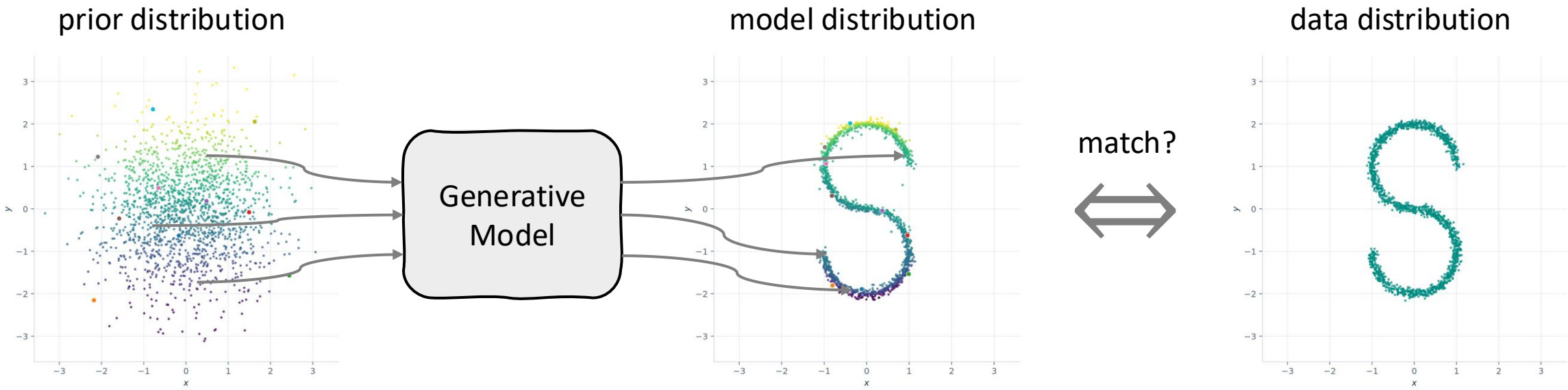


data distribution



- **Task**: transform a prior distribution into the data distribution
- **Prior distribution**: easy to sample from, e.g., standard Gaussian
- **Data distribution**: usually concentrated near a low-dim. manifold in a high-dim. space

Generative Modeling

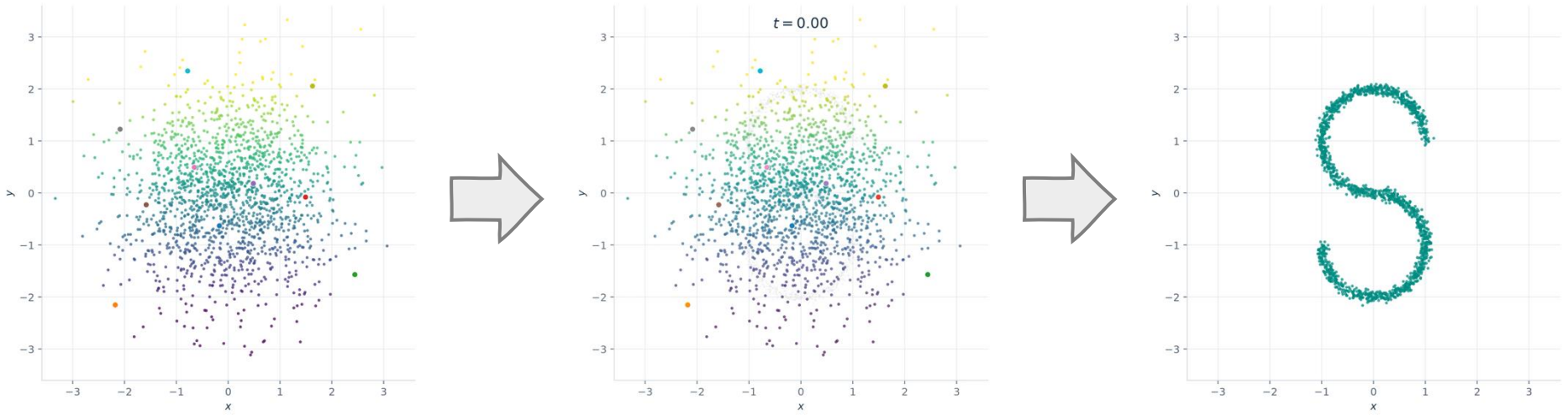


- **Model**: single-step or multi-step (iterative)
- **Model**: determinisic or stochastic
- **Goal**: distribution matching

} Many types of generative models

Continuous Normalizing Flow

prior distribution

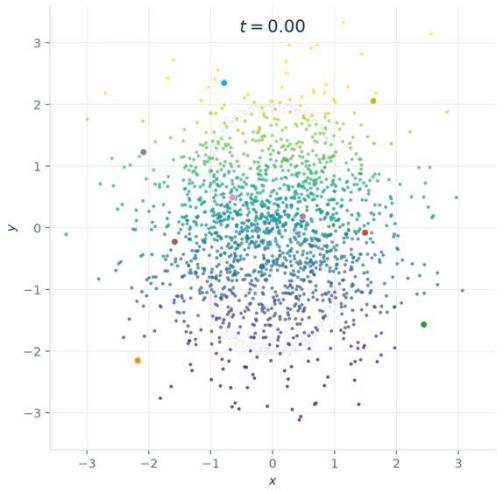


data distribution

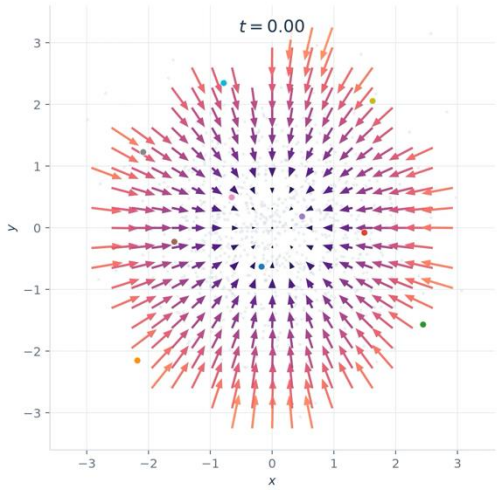
- **Intuition**: each particle moves along a direction at each timestep
- **Continuous**: timestep is continuous, usually [0,1]
- **Normalizing**: the total probability mass remains equal to “1”

Continuous Normalizing Flow

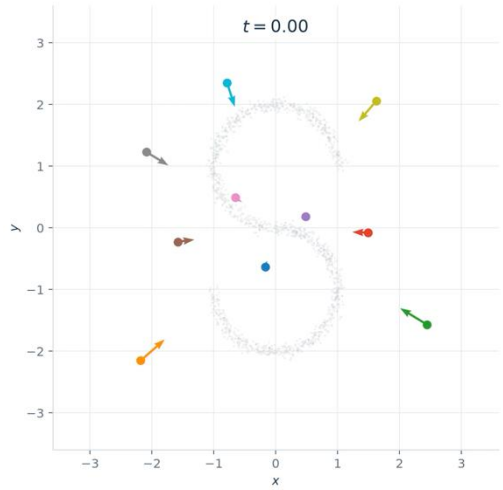
probability density paths



velocity field



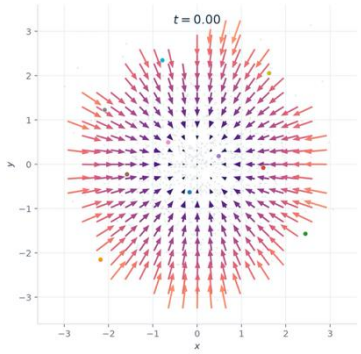
flow (trajectory)



- **Probability density paths**: distribution of particles at time t — $p: [0,1] \times \mathbb{R}^d \to \mathbb{R}_{>0}, \quad (t, \mathbf{x}) \mapsto p_t(\mathbf{x})$
- **Velocity field**: speed and direction at position x at time t — $\mathbf{u}: [0,1] \times \mathbb{R}^d \to \mathbb{R}^d, \quad (t, \mathbf{x}) \mapsto \mathbf{u}_t(\mathbf{x})$
- **Flow**: position at time t from starting point x — $\boldsymbol{\phi}: [0,1] \times \mathbb{R}^d \to \mathbb{R}^d, \quad (t, \mathbf{x}) \mapsto \boldsymbol{\phi}_t(\mathbf{x})$

Continuous Normalizing Flow

velocity field



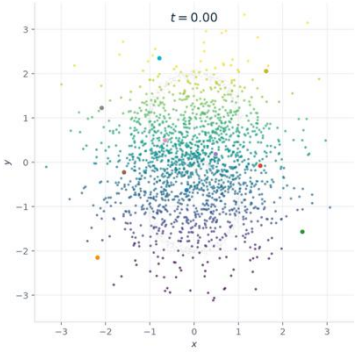
Continuity Equation

$$\frac{\partial}{\partial t} p_t(\mathbf{x}) + \nabla \cdot \left(p_t(\mathbf{x}) \mathbf{u}_t(\mathbf{x}) \right) = 0$$

Ordinary Differential Equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{\phi}_t(\mathbf{x}) = \mathbf{u}_t\left(\boldsymbol{\phi}_t(\mathbf{x})\right), \quad \boldsymbol{\phi}_0(\mathbf{x}) = \mathbf{x}$$

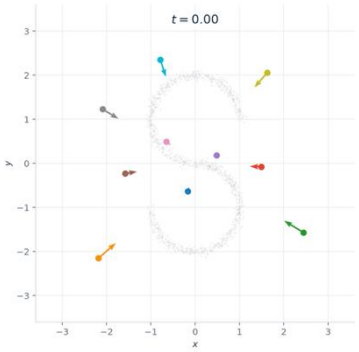
probability density paths



Push-Forward Equation

$$p_t(\mathbf{x}) = p_0\left(\boldsymbol{\phi}_t^{-1}(\mathbf{x})\right) \left| \det \left[\frac{\partial}{\partial \mathbf{x}} \boldsymbol{\phi}_t^{-1}(\mathbf{x}) \right] \right|$$

flow (trajectory)



Flow Matching

- **Training**: velocity field regression

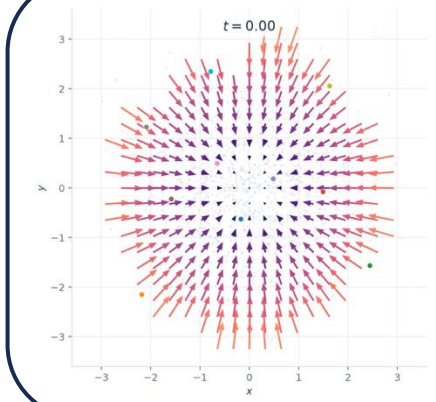
$$\mathcal{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x})\|^2]$$

velocity prediction model

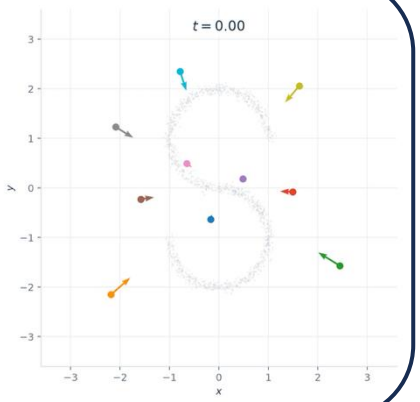
ground-truth velocity field

We don't know the ground-truth velocity field!

- **Sampling**: ODE integration from noise to data



$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}_t(\mathbf{x}) = \mathbf{u}_t(\boldsymbol{\phi}_t(\mathbf{x})),$$
$$\boldsymbol{\phi}_0(\mathbf{x}) = \mathbf{x}$$



$\Rightarrow$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}_t = \mathbf{v}_t(\mathbf{x}_t;\theta^*), \quad \mathbf{x}_0 \sim p_{\mathrm{prior}}$$

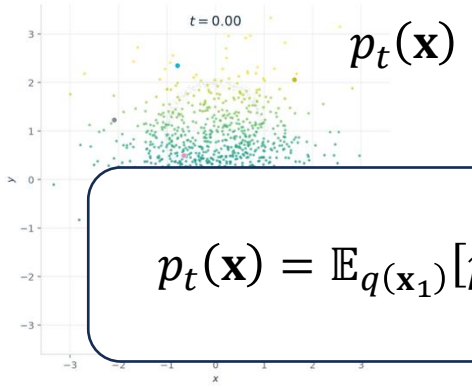
Solve the ODE from $t = 0$ to $t = 1$

Conditional Flow Matching

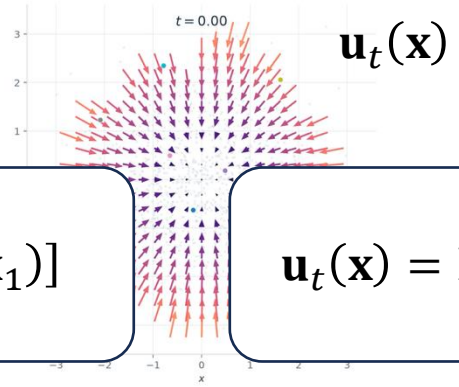
- Conditioning on a data point $\mathbf{x}_1$

probability density paths | velocity field | flow (trajectory)

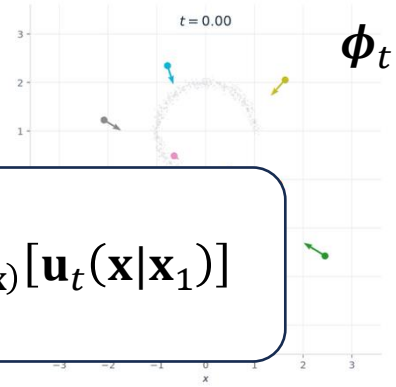
marginal



$p_t(\mathbf{x})$



$\mathbf{u}_t(\mathbf{x})$

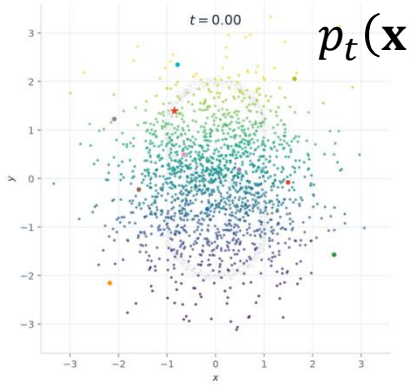


$\boldsymbol{\phi}_t(\mathbf{x})$

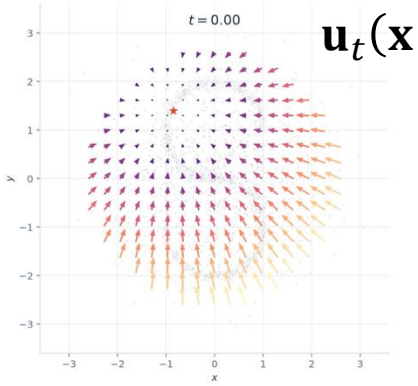
$$p_t(\mathbf{x}) = \mathbb{E}_{q(\mathbf{x}_1)}[p_t(\mathbf{x}|\mathbf{x}_1)]$$

$$\mathbf{u}_t(\mathbf{x}) = \mathbb{E}_{p(\mathbf{x}_1|\mathbf{x})}[\mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)]$$

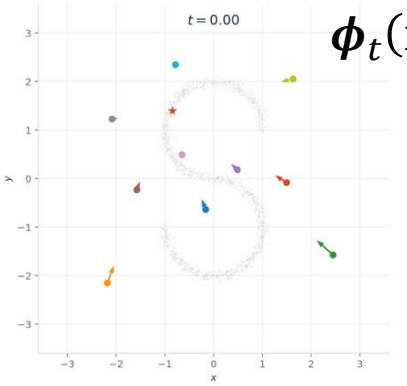
conditional



$p_t(\mathbf{x}|\mathbf{x}_1)$



$\mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)$



$\boldsymbol{\phi}_t(\mathbf{x}|\mathbf{x}_1)$

Conditional Flow Matching

Flow Matching same gradient **Conditional Flow Matching**

$$\mathcal{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x})\|^2] \quad \Leftrightarrow \quad \mathcal{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{t,q(\mathbf{x}_1),p_t(\mathbf{x}|\mathbf{x}_1)}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\|^2]$$

unknown marginal velocity — known conditional velocity

$$\begin{aligned}
\mathcal{L}_{\mathrm{FM}}(\theta) &= \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta)\|^2] + \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{u}_t(\mathbf{x})\|^2] - 2\mathbb{E}_{t,p_t(\mathbf{x})}[\langle \mathbf{v}_t(\mathbf{x};\theta), \mathbf{u}_t(\mathbf{x})\rangle] \\
&\cong \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta)\|^2] - 2\int \left\langle \mathbf{v}_t(\mathbf{x};\theta), \int \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1) p_t(\mathbf{x}|\mathbf{x}_1) q(\mathbf{x}_1) \mathrm{d}\mathbf{x}_1 \right\rangle \mathrm{d}\mathbf{x} \\
&= \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta)\|^2] - 2\iint \langle \mathbf{v}_t(\mathbf{x};\theta), \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\rangle\, p_t(\mathbf{x}|\mathbf{x}_1) q(\mathbf{x}_1) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x} \\
&= \mathbb{E}_{t,q(\mathbf{x}_1),p_t(\mathbf{x}|\mathbf{x}_1)}[\|\mathbf{v}_t(\mathbf{x};\theta)\|^2 - 2\langle \mathbf{v}_t(\mathbf{x};\theta), \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\rangle] \\
&\cong \mathbb{E}_{t,q(\mathbf{x}_1),p_t(\mathbf{x}|\mathbf{x}_1)}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\|^2] = \mathcal{L}_{\mathrm{CFM}}(\theta)
\end{aligned}$$

Conditional Flow Matching

Flow Matching

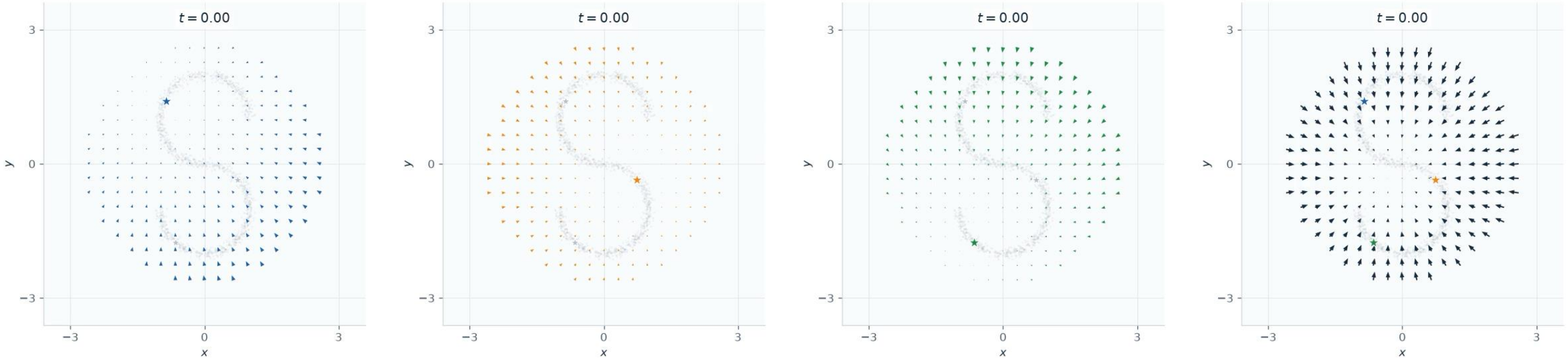
$$\mathcal{L}_{\mathrm{FM}}(\theta) = \mathbb{E}_{t,p_t(\mathbf{x})}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x})\|^2]$$

same gradient

$\Leftrightarrow$

Conditional Flow Matching

$$\mathcal{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{t,q(\mathbf{x}_1),p_t(\mathbf{x}|\mathbf{x}_1)}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\|^2]$$



Conditional Flow Matching

- Rewrite the loss with the conditional flow

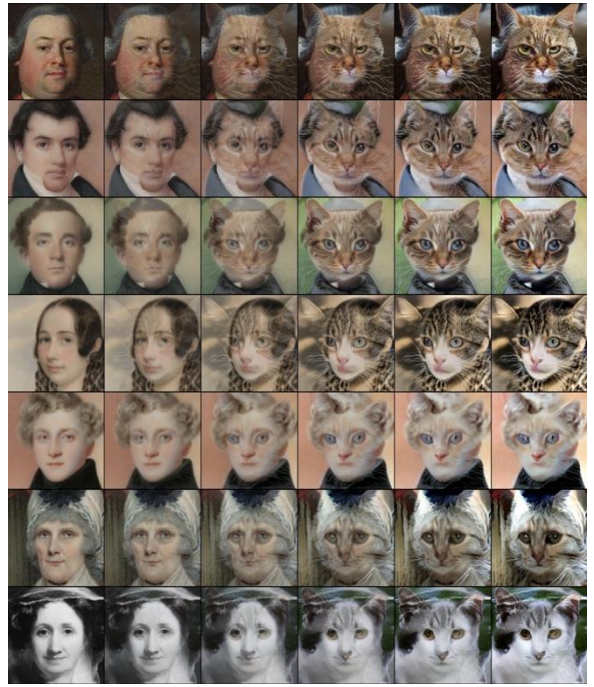
$$
\begin{aligned}
\mathcal{L}_{\mathrm{CFM}}(\theta) &= \mathbb{E}_{t,q(\mathbf{x}_1),p_t(\mathbf{x}|\mathbf{x}_1)}[\|\mathbf{v}_t(\mathbf{x};\theta) - \mathbf{u}_t(\mathbf{x}|\mathbf{x}_1)\|^2] \\
&= \mathbb{E}_{t,q(\mathbf{x}_1),p(\mathbf{x}_0)}[\|\mathbf{v}_t(\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1);\theta) - \mathbf{u}_t(\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1)|\mathbf{x}_1)\|^2] \\
&= \mathbb{E}_{t,q(\mathbf{x}_1),p(\mathbf{x}_0)}\left[\left\|\mathbf{v}_t(\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1);\theta) - \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1)\right\|^2\right]
\end{aligned}
$$

- One particular instantiation
  - Set the prior distribution to **standard Gaussian**: $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{0}, \mathbf{I})$
  - Set the conditional flow to **linear interpolation**: $\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = (1-t)\mathbf{x}_0 + t\mathbf{x}_1$
  - Then, the regression target becomes: $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = \mathbf{x}_1 - \mathbf{x}_0$

$$
\mathcal{L}_{\mathrm{CFM}}(\theta) = \mathbb{E}_{t,q(\mathbf{x}_1),p(\mathbf{x}_0)}\left[\left\|\mathbf{v}_t\big((1-t)\mathbf{x}_0 + t\mathbf{x}_1\big);\theta\big) - (\mathbf{x}_1 - \mathbf{x}_0)\right\|^2\right]
$$

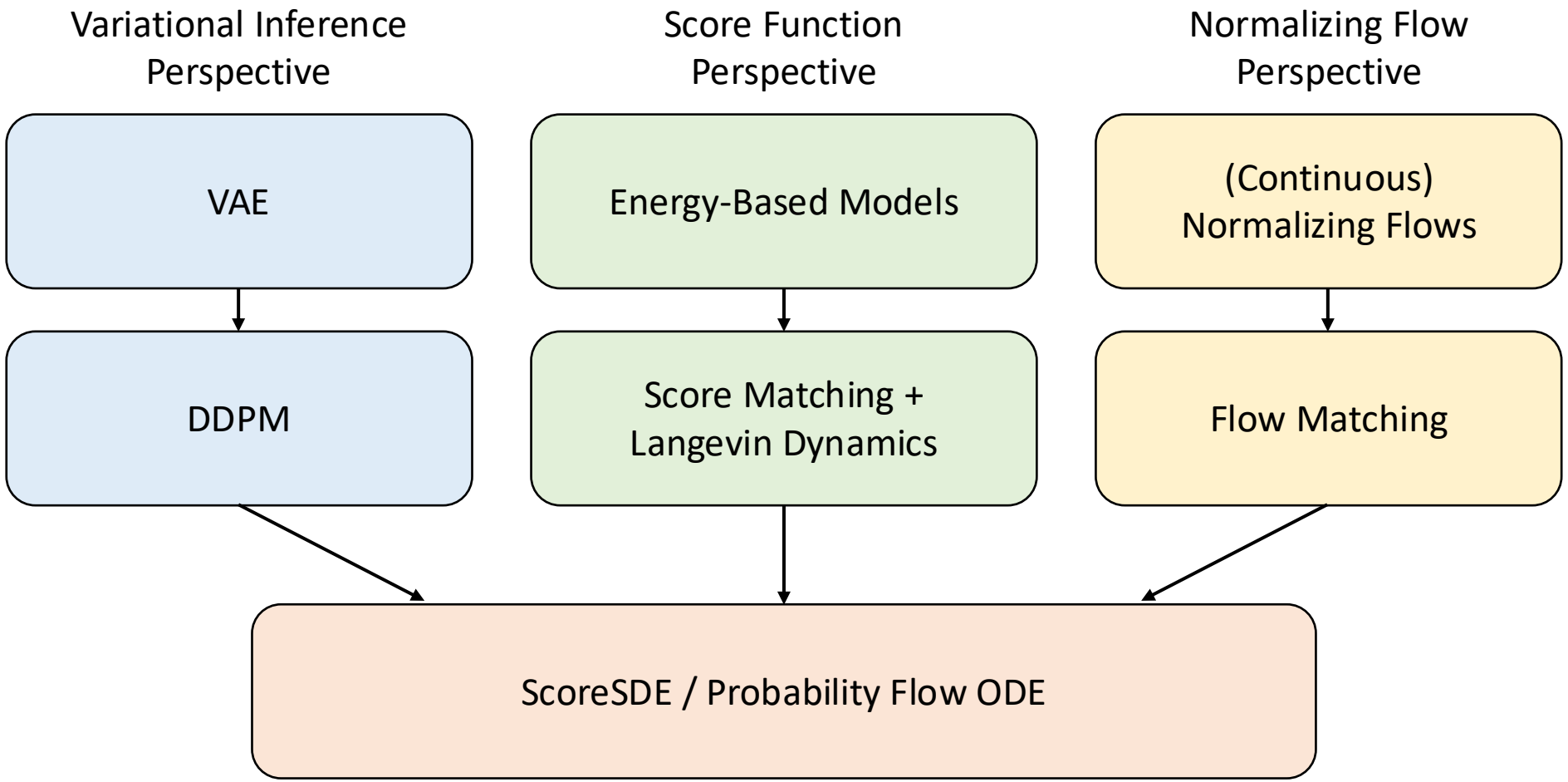
Extension

- One particular instantiation
  - Set the prior distribution to **standard Gaussian**: $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{0}, \mathbf{I})$
  - Set the conditional flow to **linear interpolation**: $\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = (1-t)\mathbf{x}_0 + t\mathbf{x}_1$
  - Then, the regression target becomes: $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = \mathbf{x}_1 - \mathbf{x}_0$
- Beyond Gaussian prior distribution
  - Any source distribution $p(\mathbf{x}_0)$ and any target distribution $p(\mathbf{x}_1)$
  - Application: image-to-image translation
- Beyond linear interpolation
  - Set the conditional flow to: $\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = \alpha_t\mathbf{x}_0 + \beta_t\mathbf{x}_1$
  - Then, the regression target becomes: $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\phi}_t(\mathbf{x}_0|\mathbf{x}_1) = \dot{\alpha}_t\mathbf{x}_0 + \dot{\beta}_t\mathbf{x}_1$



Credit: Liu et.al. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow. ICLR 2023

Relation to Diffusion & Score-Based Gen.



Reference: Lai et.al. The Principles of Diffusion Models: From Origins to Advances. arXiv:2510.21890

Resources

- **Papers**
  - Lipman et.al. Flow Matching for Generative Modeling.
  - Liu et.al. Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow.
  - Albergo et.al. Building Normalizing Flows with Stochastic Interpolants.
- **Blogs / Tutorials**
  - Yang Song. Generative Modeling by Estimating Gradients of the Data Distribution.
  - Lilian Weng. What are Diffusion Models?
  - Lipman et.al. Flow Matching Guide and Code.
  - Lai et.al. The Principles of Diffusion Models.
- **Courses**
  - Stanford CS236: Deep Generative Models
  - CMU 10-799 Diffusion & Flow Matching
  - MIT 6.S184: Introduction to Flow Matching and Diffusion Models

Thanks