Lecture 2

Flow Matching

MIT IAP 2026 | Jan 22, 2025

Peter Holderrieth and Ron Shprints

Sponsor: Tommi Jaakkola

Reminder: Flow and Diffusion Models

Flow Model

Initialize: $X_0 \sim p_{\text{init}},$

ODE: $\mathrm{d}X_t = u_t^\theta(X_t)\mathrm{d}t$

E.g. Gaussian

Neural network vector field

Diffusion coeff.

Diffusion Model

Initialize: $X_0 \sim p_{\text{init}},$

SDE: $\mathrm{d}X_t = u_t^\theta(X_t)\mathrm{d}t + \sigma_t \mathrm{d}W_t$

To get samples, simulate ODE/SDE from t=0 to t=1 and return X_1

Next step: Training a flow model

Without training, the model produces “non-sense”→ **We need to train** u_t^θ

Training = Finding parameters θ **such that**

$X_0 \sim p_{\text{init}}, \quad \mathrm{d}X_t = u_t^\theta(X_t)\mathrm{d}t$ → Implies → $X_1 \sim p_{\text{data}}$

Start with initial distribution

Follow along the vector field

The distribution of the final point = data distribution

Goal of lecture 2 (today):

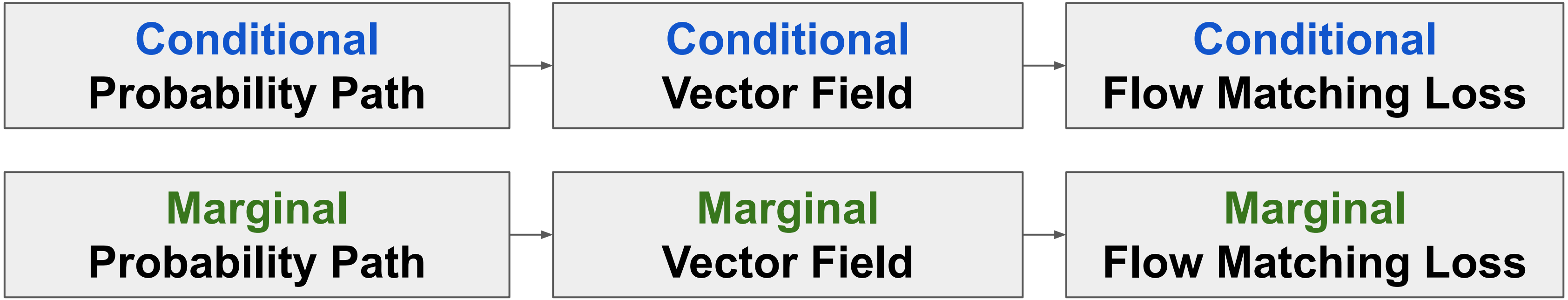
Derive Flow Matching (training algorithm for flow models)

Section 2:

Flow Matching

Goal: Derive a training algorithm for flow models

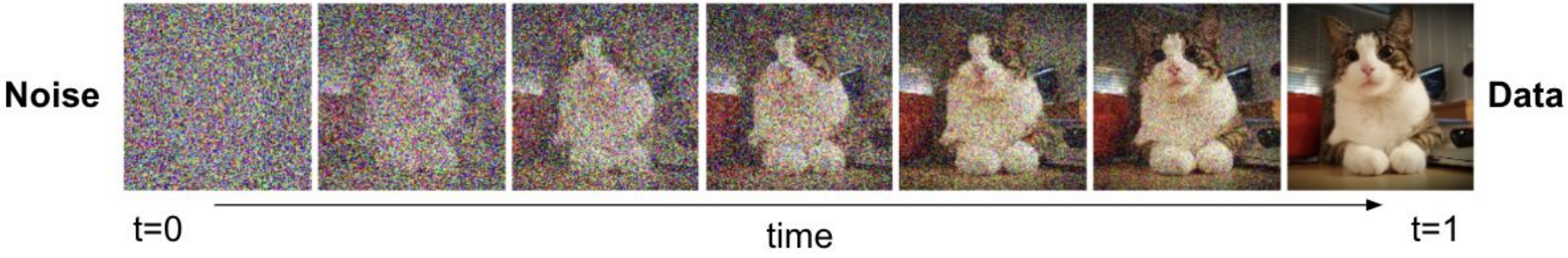
The Flow Matching Matrix

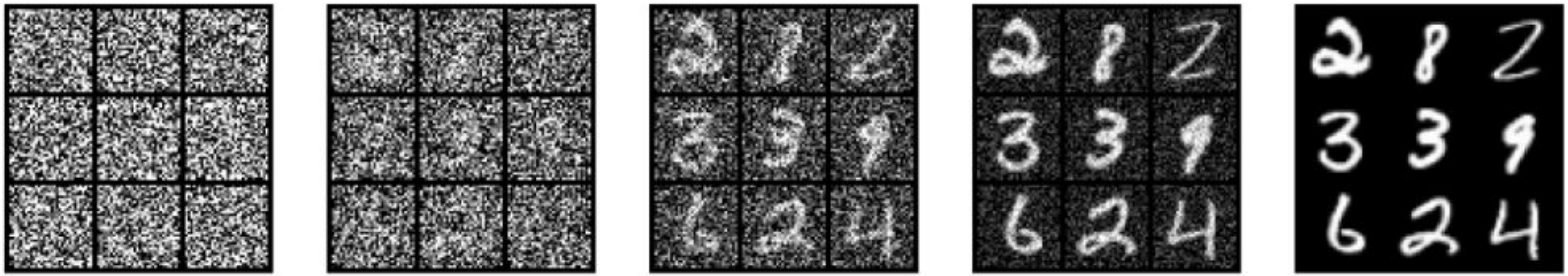


“Conditional” = “Per single data point”

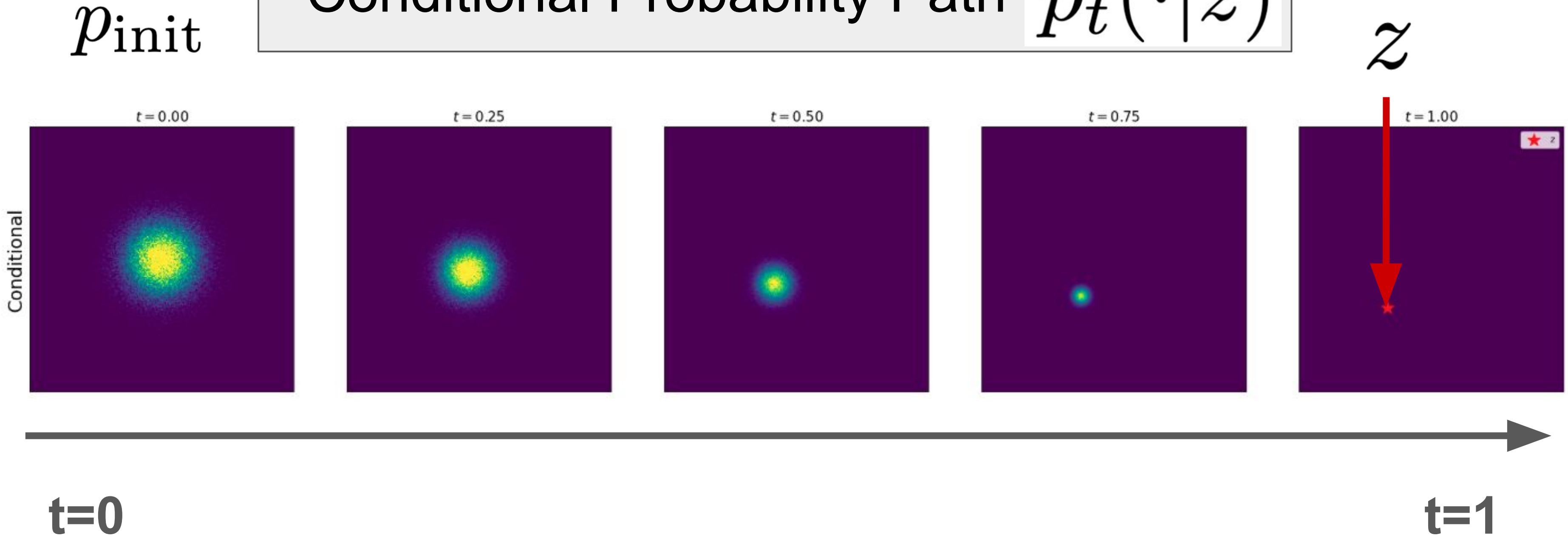
“Marginal” = “Across distribution of data points”

Probability Paths: The Path from Noise to Data

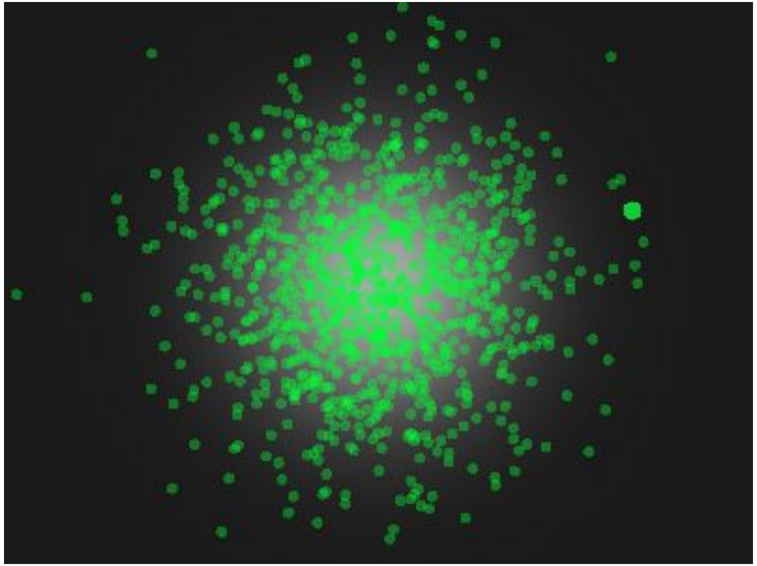




Conditional Probability Path $p_t(\cdot|z)$

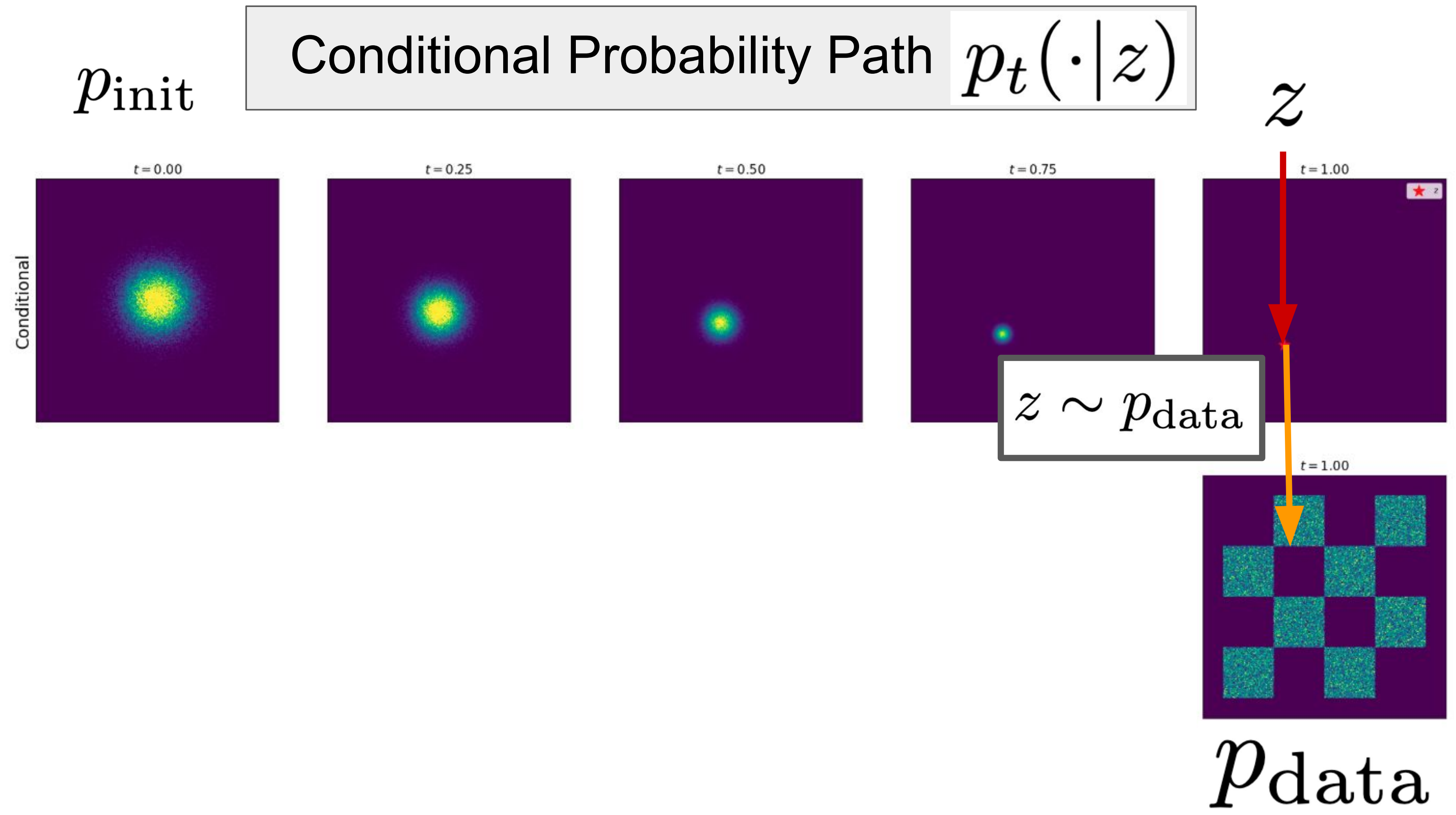


Plotting samples from a conditional probability path over time

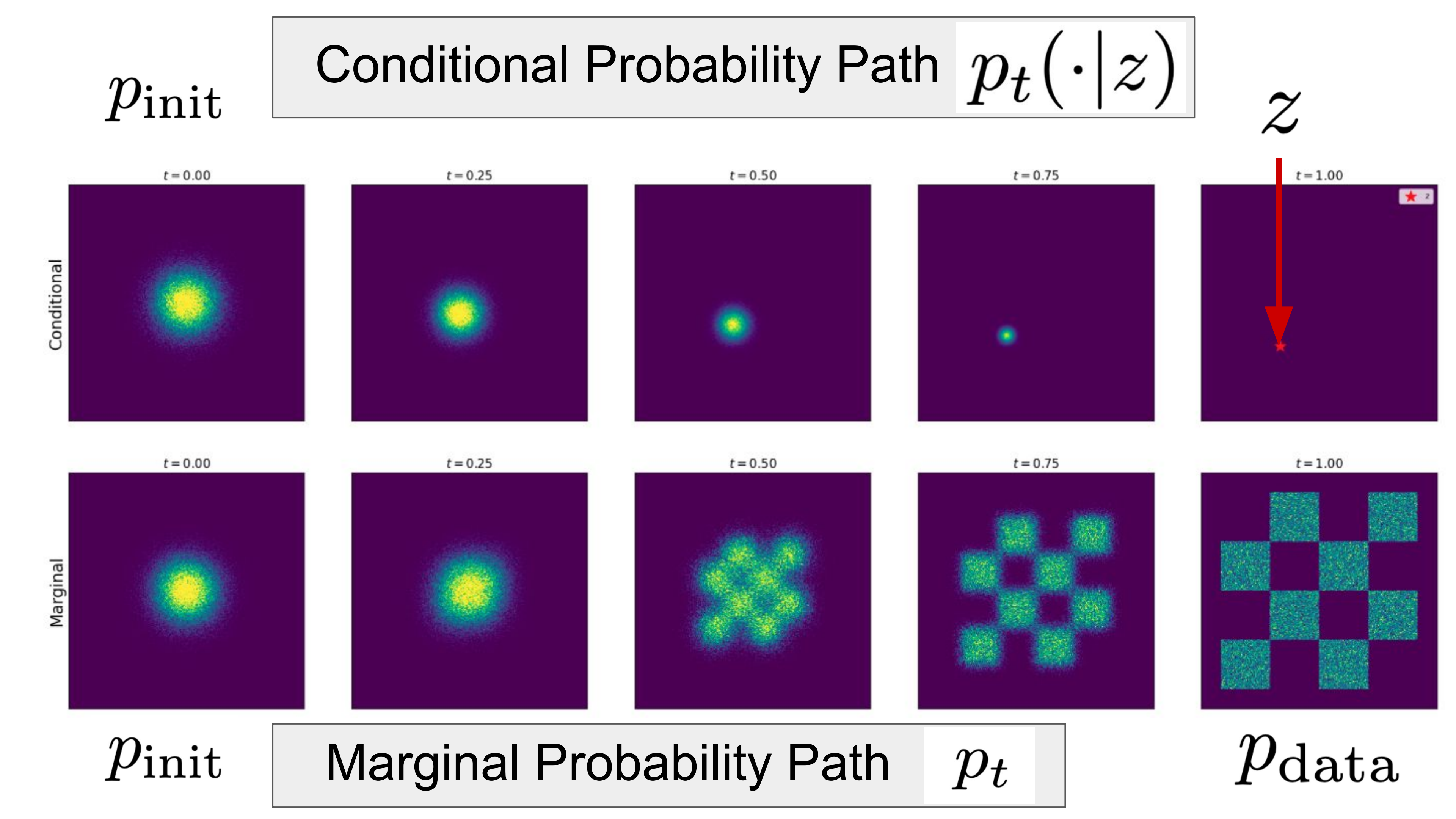


*Note: A probability path only specifies the marginals (each snapshot). It says **nothing about the evolution of a single particle** in time (no dynamics).*

Source: https://mariogemoll.com/flow-matching



p_{init}
Conditional Probability Path $p_t(\cdot|z)$
z
$t=0.00$
$t=0.25$
$t=0.50$
$t=0.75$
$t=1.00$
Conditional
$z \sim p_{\text{data}}$
$t=1.00$
p_{data}



p_{init}
Conditional Probability Path $p_t(\cdot|z)$
z
t = 0.00
t = 0.25
t = 0.50
t = 0.75
t = 1.00
Conditional
Marginal
p_{init}
Marginal Probability Path p_t
p_{data}

Conditional Prob. Path

	Notation	Key property	Gaussian example
Conditional Probability Path	$p_t(\cdot\|z)$	Interpolates p_{init} and a data point z	$\mathcal{N}(\alpha_t z, \beta_t^2 I_d)$
Conditional Vector Field			

Marginal Prob. Path

	Notation	Key property	Formula
Marginal Probability Path	p_t	Interpolates p_{init} and p_{data}	$\int p_t(x\|z)p_{\text{data}}(z)\mathrm{d}z$
Marginal Vector Field			

Example - Conditional Vector Field for Gaussian

$$u_t^{\text{target}}(x|z) = \left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z + \frac{\dot{\beta}_t}{\beta_t} x$$

Proof Sketch:

Step 1: By checking ODE, show that the flow of the vector field is given by

$$\psi_t^{\text{target}}(x_0|z) = \alpha_t z + \beta_t x_0$$

Step 2: If $X_0 = x_0 \sim \mathcal{N}(0, I_d)$ is random, then we know that then:

$$X_t = \psi_t(X_0|z) = \alpha_t z + \beta_t X_0 \sim \mathcal{N}(\alpha_t z, \beta_t^2 I_d) = p_t(\cdot|z)$$

Gaussian Conditional Probability Path And Conditional Vector Field

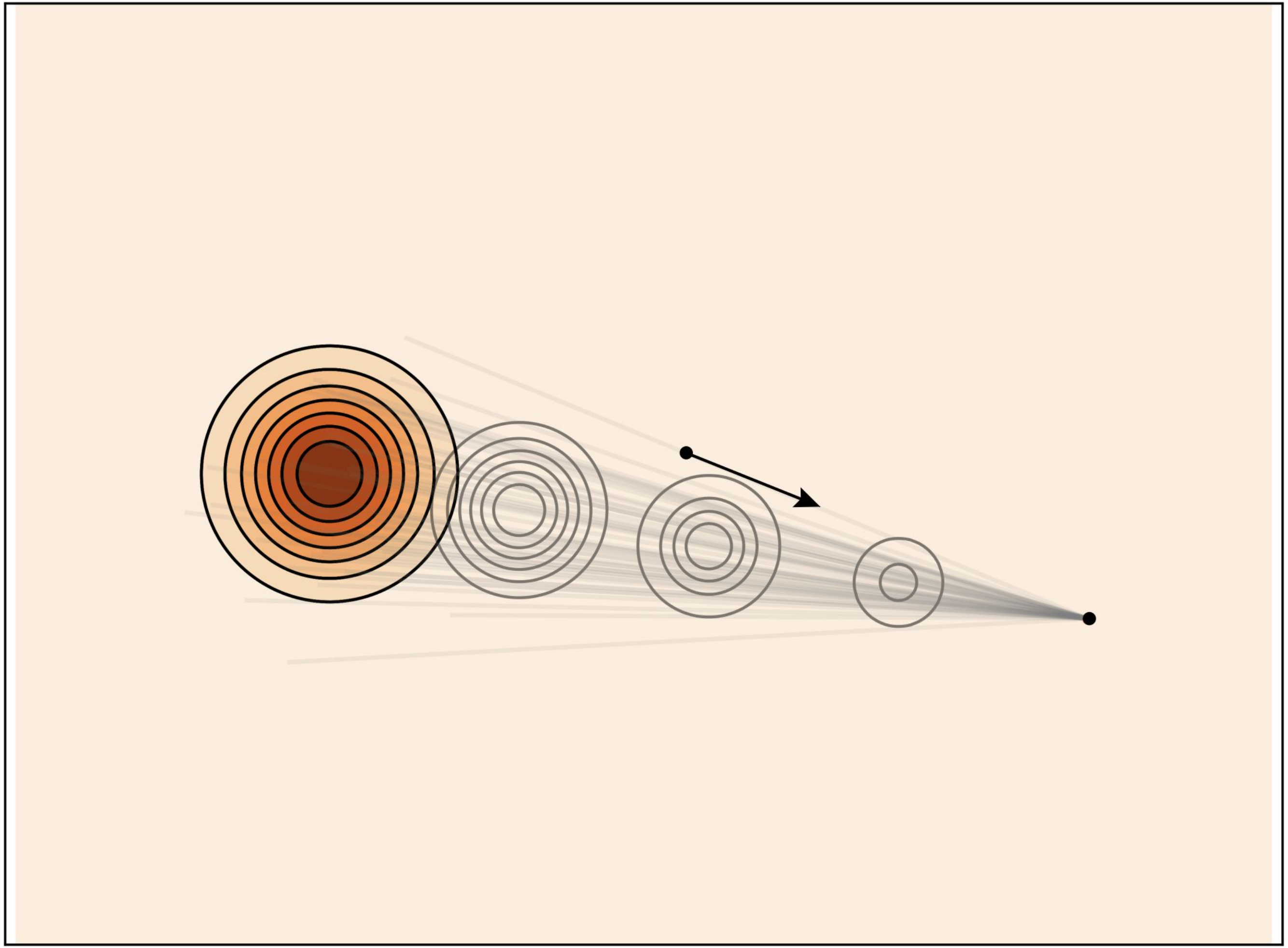
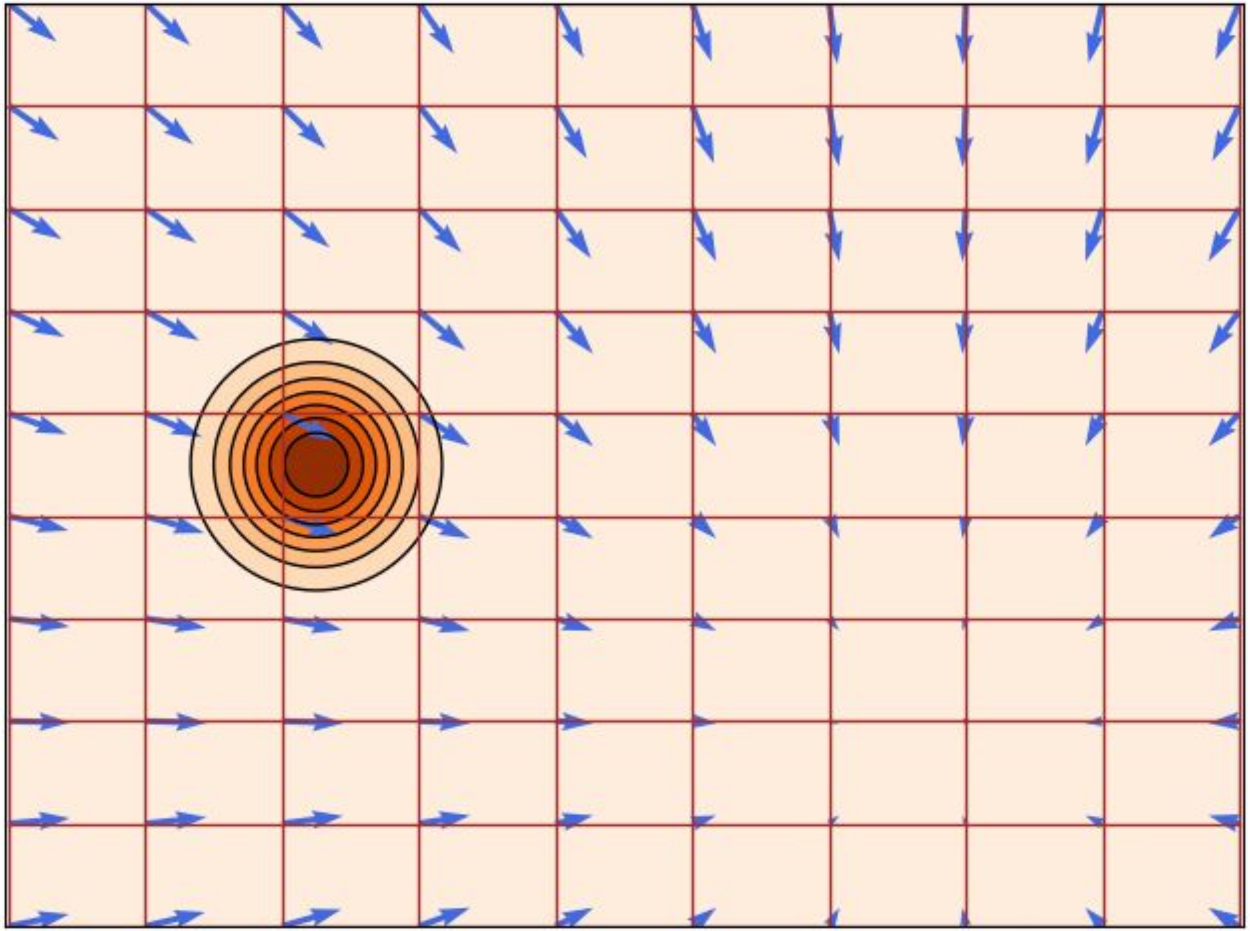


Figure credit: Yaron Lipman

Simulating ODE with Conditional Vector Field for Conditional Probability Path



NOTE: This is an animated gif and is static in a PDF

Figure credit:
Yaron Lipman

Toy example

NOTE: This is an animated gif and is static in a PDF

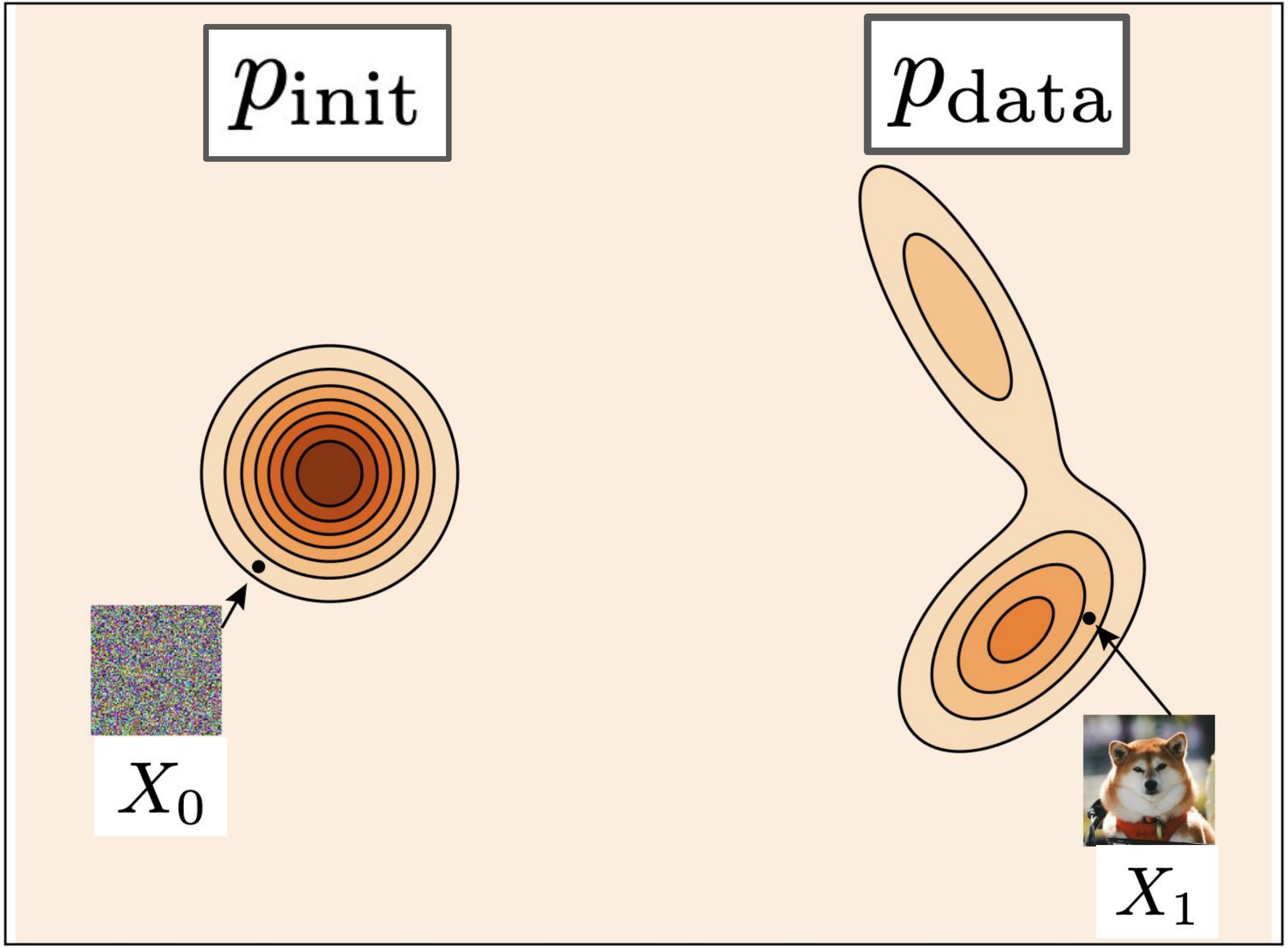


Figure credit: Yaron Lipman

Simulating ODE with Marginal Vector Field for Gaussian Probability Path

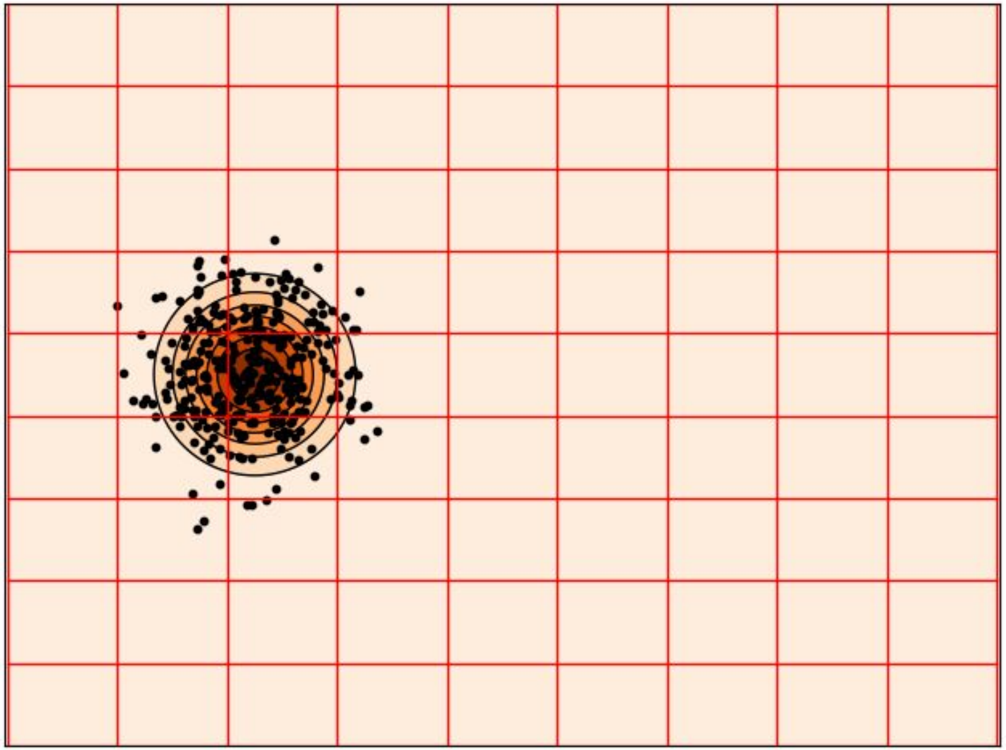
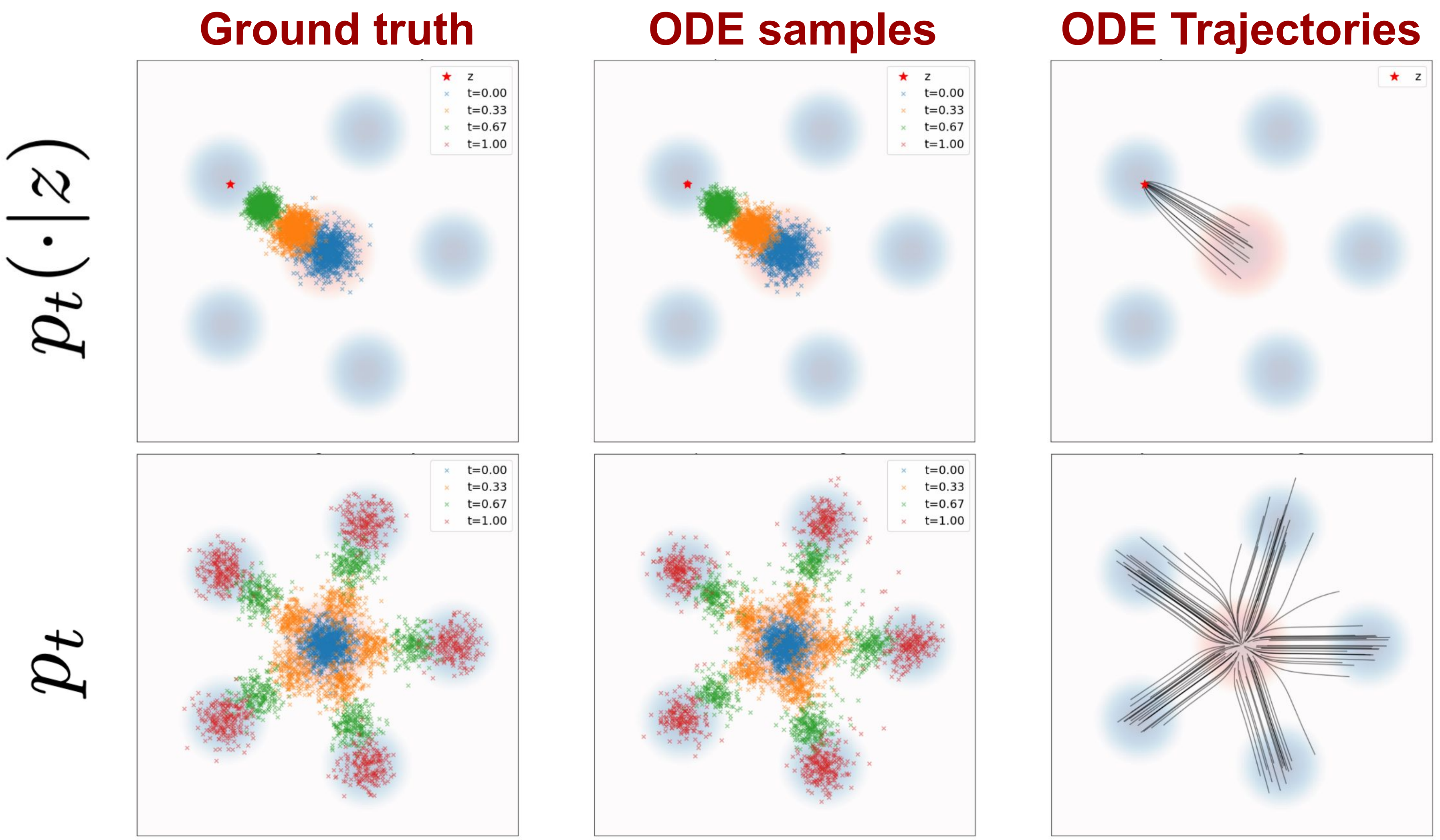


Figure credit: Yaron Lipman


Ground truth
ODE samples
ODE Trajectories
$p_t(\cdot|z)$
p_t
z
t=0.00
t=0.33
t=0.67
t=1.00

Continuity Equation

Randomly initialized ODE

Given: $$X_0 \sim p_{\text{init}}, \qquad \frac{\mathrm{d}}{\mathrm{d}t} X_t = u_t(X_t)$$

Follow probability path:

$$X_t \sim p_t \quad (0 \leq t \leq 1)$$

Marginals are p_t

equivalent

Continuity equation holds

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t(x) = -\mathrm{div}(p_t u_t)(x)$$

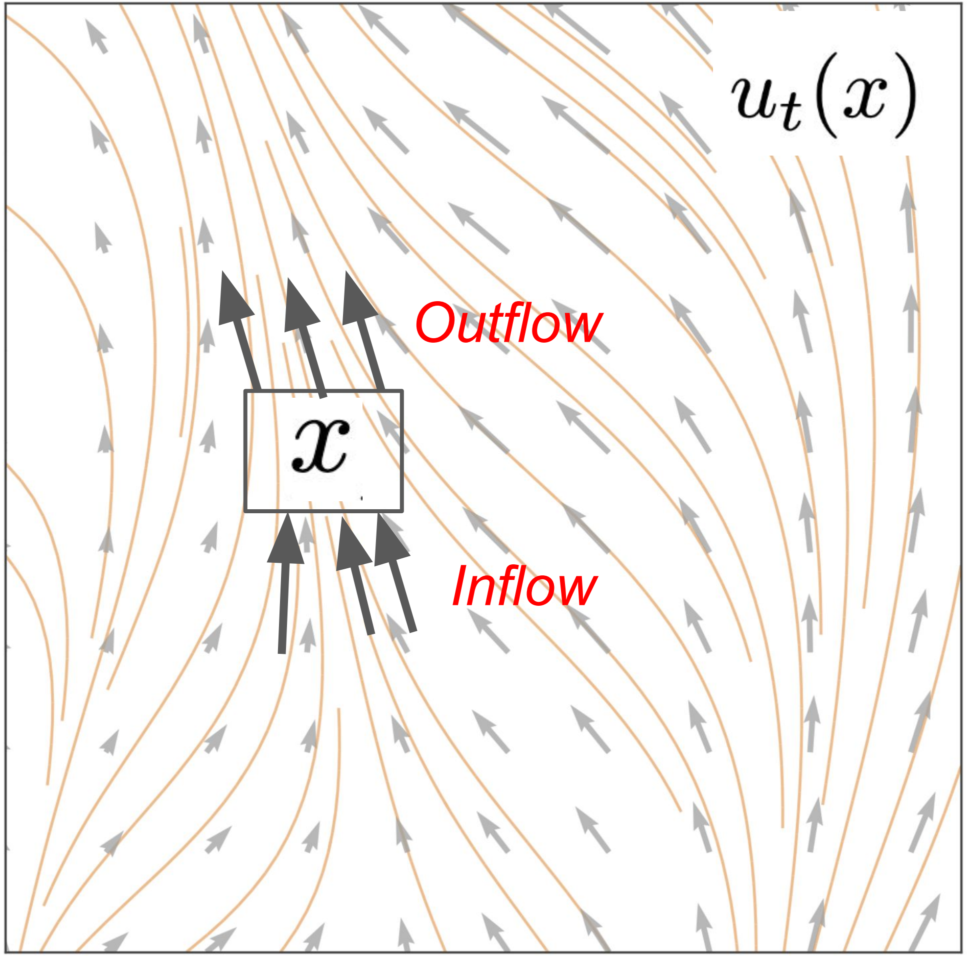
PDE holds

Continuity Equation

$$\frac{\mathrm{d}}{\mathrm{d}t} p_t(x) = -\mathrm{div}(p_t u_t)(x)$$

Change of probability mass at x

Outflow - inflow of probability mass from u



Conditional Prob. Path, Vector Field, and Score

	Notation	Key property	Gaussian example
Conditional Probability Path	$p_t(\cdot \mid z)$	Interpolates p_{init} and a data point z	$\mathcal{N}(\alpha_t z, \beta_t^2 I_d)$
Conditional Vector Field	$u_t^{\text{target}}(x \mid z)$	ODE follows conditional path	$\left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z + \frac{\dot{\beta}_t}{\beta_t} x$

Marginal Prob. Path, Vector Field, and Score

	Notation	Key property	Formula
Marginal Probability Path	p_t	Interpolates p_{init} and p_{data}	$\int p_t(x\|z)p_{\mathrm{data}}(z)\mathrm{d}z$
Marginal Vector Field	$u_t^{\mathrm{target}}(x)$	ODE follows marginal path	$\int u_t^{\mathrm{target}}(x\|z)\frac{p_t(x\|z)p_{\mathrm{data}}(z)}{p_t(x)}\mathrm{d}z$

Example marginal vector field - Meta MovieGen

These videos are generated by simulating the ODE with the (learnt) marginal vector field

Algorithm 3 Flow Matching Training Procedure (General)

Require: A dataset of samples $z \sim p_{\text{data}}$, neural network u_t^θ

for each mini-batch of data **do**

  Sample a data example z from the dataset.

  Sample a random time $t \sim \text{Unif}_{[0,1]}$.

  Sample $x \sim p_t(\cdot|z)$

  Compute loss

$$\mathcal{L}(\theta) = \|u_t^\theta(x) - u_t^{\text{target}}(x|z)\|^2$$

  Update the model parameters θ via gradient descent on $\mathcal{L}(\theta)$

end for

Conditional Flow Matching for Gaussian probability path

Prob. path $\mathcal{N}(\alpha_t z, \beta_t^2 I_d)$ Conditional VF $u_t^{\text{target}}(x|z) = \left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z + \frac{\dot{\beta}_t}{\beta_t} x$

Noise Sampling $x \sim p_t(\cdot|z) \quad \Leftrightarrow \quad x = \alpha_t z + \beta_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, I_d)$

Plugging in Noise Sampling into CFM Loss results in:

$$
\begin{aligned}
&L_{\text{CFM}}(\theta) \\
=&\mathbb{E}_{t\sim\text{Unif},\, z\sim p_{\text{data}},\, x\sim p_t(\cdot|z)}\left[\|u_t^\theta(x) - u_t^{\text{target}}(x|z)\|^2\right] \\
=&\mathbb{E}_{t\sim\text{Unif},\, z\sim p_{\text{data}},\, \epsilon\sim\mathcal{N}(0,I_d)}\left[\|u_t^\theta(\alpha_t z + \beta_t\epsilon) - u_t^{\text{target}}(\alpha_t z + \beta_t\epsilon|z)\|^2\right] \\
=&\mathbb{E}_{t\sim\text{Unif},\, z\sim p_{\text{data}},\, \epsilon\sim\mathcal{N}(0,I_d)}\left[\|u_t^\theta(\alpha_t z + \beta_t\epsilon) - (\dot{\alpha}_t z + \dot{\beta}_t\epsilon)\|^2\right]
\end{aligned}
$$

noise+data **velocity**

Straight Line Schedule

$$
\begin{aligned}
& L_{\mathrm{CFM}}(\theta) \\
=& \mathbb{E}_{t\sim\mathrm{Unif},\, z\sim p_{\mathrm{data}},\, \epsilon\sim\mathcal{N}(0,I_d)} \left[\| u_t^\theta(\alpha_t z + \beta_t \epsilon) - (\dot{\alpha}_t z + \dot{\beta}_t \epsilon) \|^2 \right] \\
=& \mathbb{E}_{t\sim\mathrm{Unif},\, z\sim p_{\mathrm{data}},\, \epsilon\sim\mathcal{N}(0,I_d)} \left[\| u_t^\theta(tz + (1-t)\epsilon) - (z - \epsilon) \|^2 \right]
\end{aligned}
$$

Linear interpolation of noise and data

Difference between noise and data

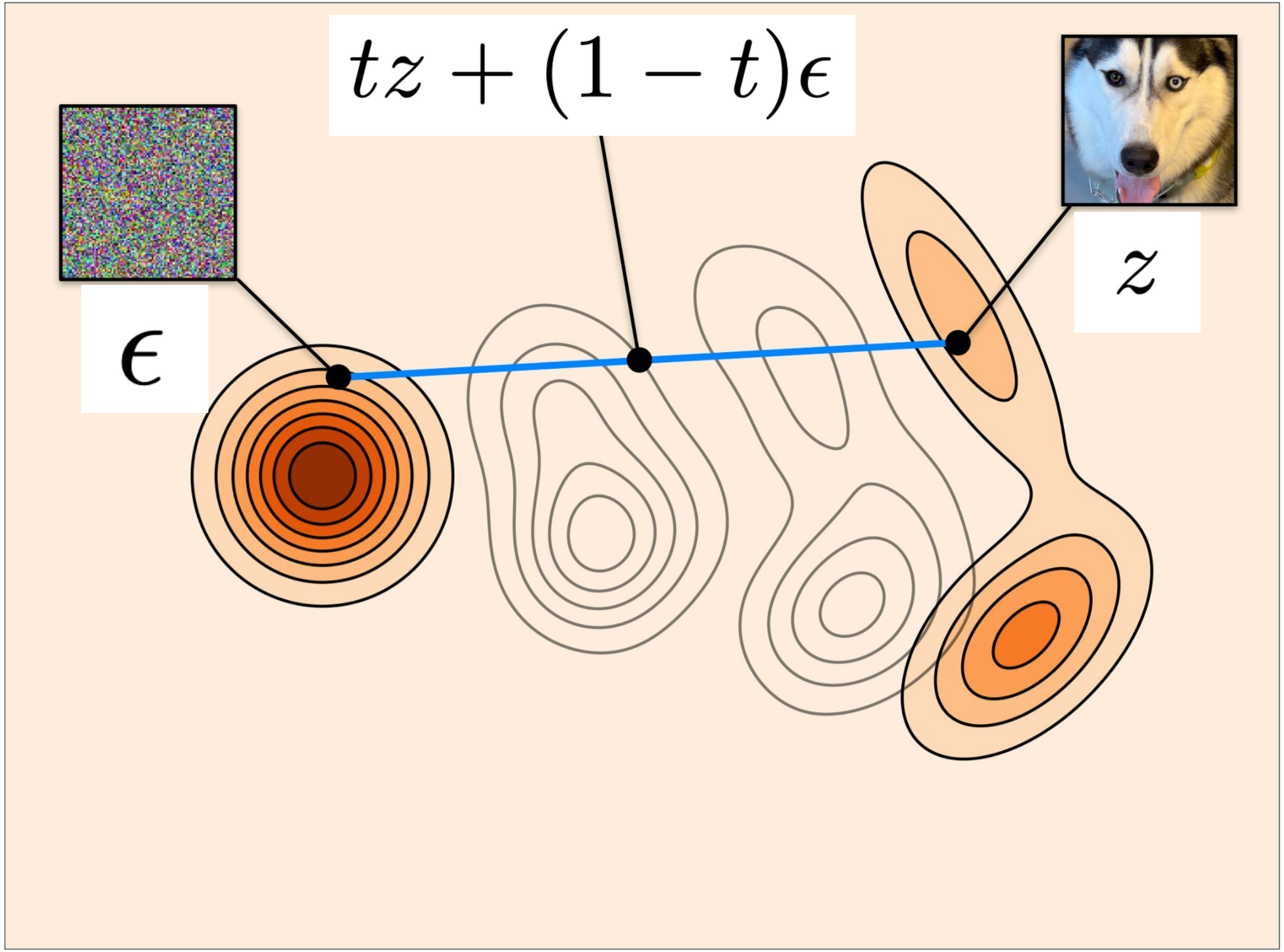


Figure credit: Yaron Lipman

Algorithm 4 Flow Matching Training for CondOT path

Require: A dataset of samples $z \sim p_{\text{data}}$, neural network u_t^θ

for each mini-batch of data **do**

  Sample a data example z from the dataset.

  Sample a random time $t \sim \text{Unif}_{[0,1]}$.

  Sample noise $\epsilon \sim \mathcal{N}(0, I_d)$

  Set $x = tz + (1-t)\epsilon$

  Compute loss

$$\mathcal{L}(\theta) = \|u_t^\theta(x) - (z - \epsilon)\|^2$$

  Update the model parameters θ via gradient descent on $\mathcal{L}(\theta)$.

end for

Example Flow Matching - Meta MovieGen

The neural network that generates these videos was trained with the algorithm in the previous slide

Example Flow Matching - Stable Diffusion 3

The neural network that generates these images was trained with the algorithm just shown

Reminder: Sampling Algorithm for Flow Model

Algorithm 1 Sampling from a Flow Model with Euler method

Require: Neural network vector field u_t^θ, number of steps n

```
Set t = 0
Set step size h = 1/n
Draw a sample X_0 ~ p_init          Random initialization!
for i = 1, ..., n − 1 do
    X_{t+h} = X_t + h u_t^θ(X_t)
    Update t ← t + h
end for
return X_1                          Return final point
```

Set $t = 0$

Set step size $h = \frac{1}{n}$

Draw a sample $X_0 \sim p_{\text{init}}$ — *Random initialization!*

for $i = 1, \ldots, n-1$ **do**

$\quad X_{t+h} = X_t + h u_t^\theta(X_t)$

$\quad$ Update $t \leftarrow t + h$

end for

return X_1 — *Return final point*

The Flow Matching Matrix

Conditional **Probability Path** →	**Conditional** **Vector Field** →	**Conditional** **Flow Matching Loss**
Marginal **Probability Path** →	**Marginal** **Vector Field** →	**Marginal** **Flow Matching Loss**
Defines distributions from noise to data	**Defines training target that we want to learn**	**Loss function that we want to minimize during traing**

Conditional Prob. Path, Vector Field, and FM Loss

	Notation	Key property	Gaussian example
Conditional Probability Path	$p_t(\cdot\|z)$	Interpolates p_{init} and a data point z	$\mathcal{N}(\alpha_t z, \beta_t^2 I_d)$
Conditional Vector Field	$u_t^{\text{target}}(x\|z)$	ODE follows conditional path	$\left(\dot{\alpha}_t - \frac{\dot{\beta}_t}{\beta_t}\alpha_t\right) z + \frac{\dot{\beta}_t}{\beta_t} x$
Conditional FM Loss	$L_{\text{CFM}}(\theta)$	Loss we minimize during training	$\mathbb{E}_{t,z,x}[\|\|u_t^\theta(x) - u_t^{\text{target}}(x\|z)\|\|^2]$

All these objects are tractable. Just analytical formulas!

Marginal Prob. Path, Vector Field, and FM Loss

	Notation	Key property	Formula
Marginal Probability Path	p_t	Interpolates p_{init} and p_{data}	$\int p_t(x\|z)p_{\text{data}}(z)\mathrm{d}z$
Marginal Vector Field	$u_t^{\text{target}}(x)$	ODE follows marginal path	$\int u_t^{\text{target}}(x\|z)\frac{p_t(x\|z)p_{\text{data}}(z)}{p_t(x)}\mathrm{d}z$
Marginal FM Loss	$L_{\text{FM}}(\theta)$	Implicitly minimized via cond FM loss	$\mathbb{E}_{t,z,x}[\|\|u_t^\theta(x) - u_t^{\text{target}}(x)\|\|^2]$

None of these objects are tractable. But we can still learn them!

Next class:

Friday (Tomorrow), 11am-12:30pm

Score matching and guidance!

E25-111 (same room)

Office hours: Tomorrow, 3pm-4:30pm in 36-156