Math Foundations for ML

10-606

Geoff Gordon

Notes and reminders

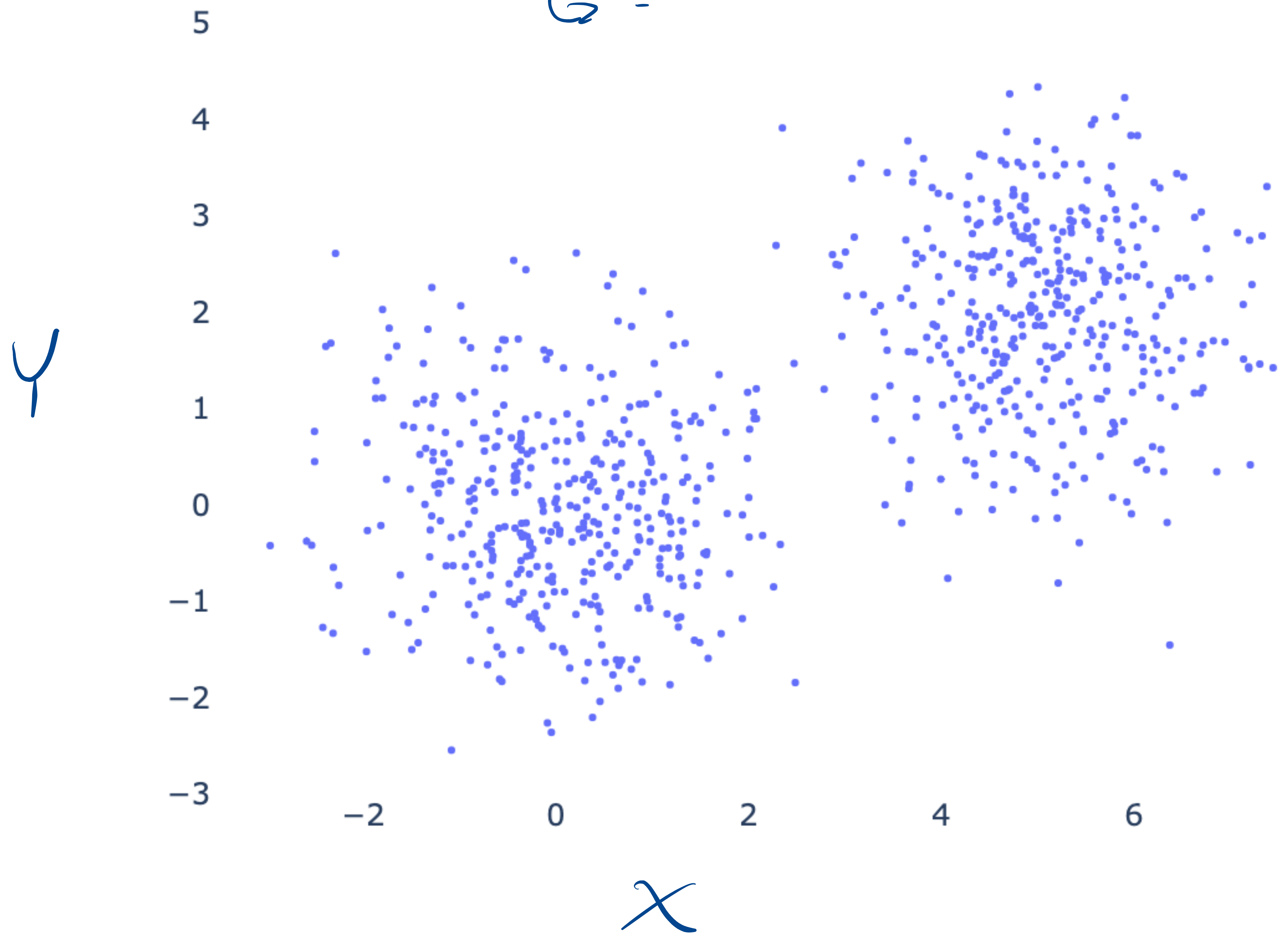
- Upcoming: Quiz 2 (Friday, different room/time)
- Please fill out FCEs!

Clusters

latent: $z = \{0,1\}$

↳ = hidden

$P(z)\,P(x|z)\,P(Y|z)$



$$P(A, B, C, D) = P(D)\,P(C \mid D)\,P(B \mid C, D)\,P(A \mid B, C, D)$$

↑ DCBA

CABD

$$= P(C)\,P(A \mid C)\,P(B \mid A, C)\,P(D \mid A, B, C)$$

$P(D \mid A, B, C)$ → maybe $= P(D \mid B)$

$P(B \mid A, C)$ → maybe $= P(B \mid A)$

model parameters Θ

data set $x_1 \dots x_T$

$y_1 \dots y_T$

$$P(\Theta)\,P(x_1, y_1 \mid \Theta)\,P(x_2, y_2 \mid \Theta) \dots P(x_T, y_T \mid \Theta)$$

(same form

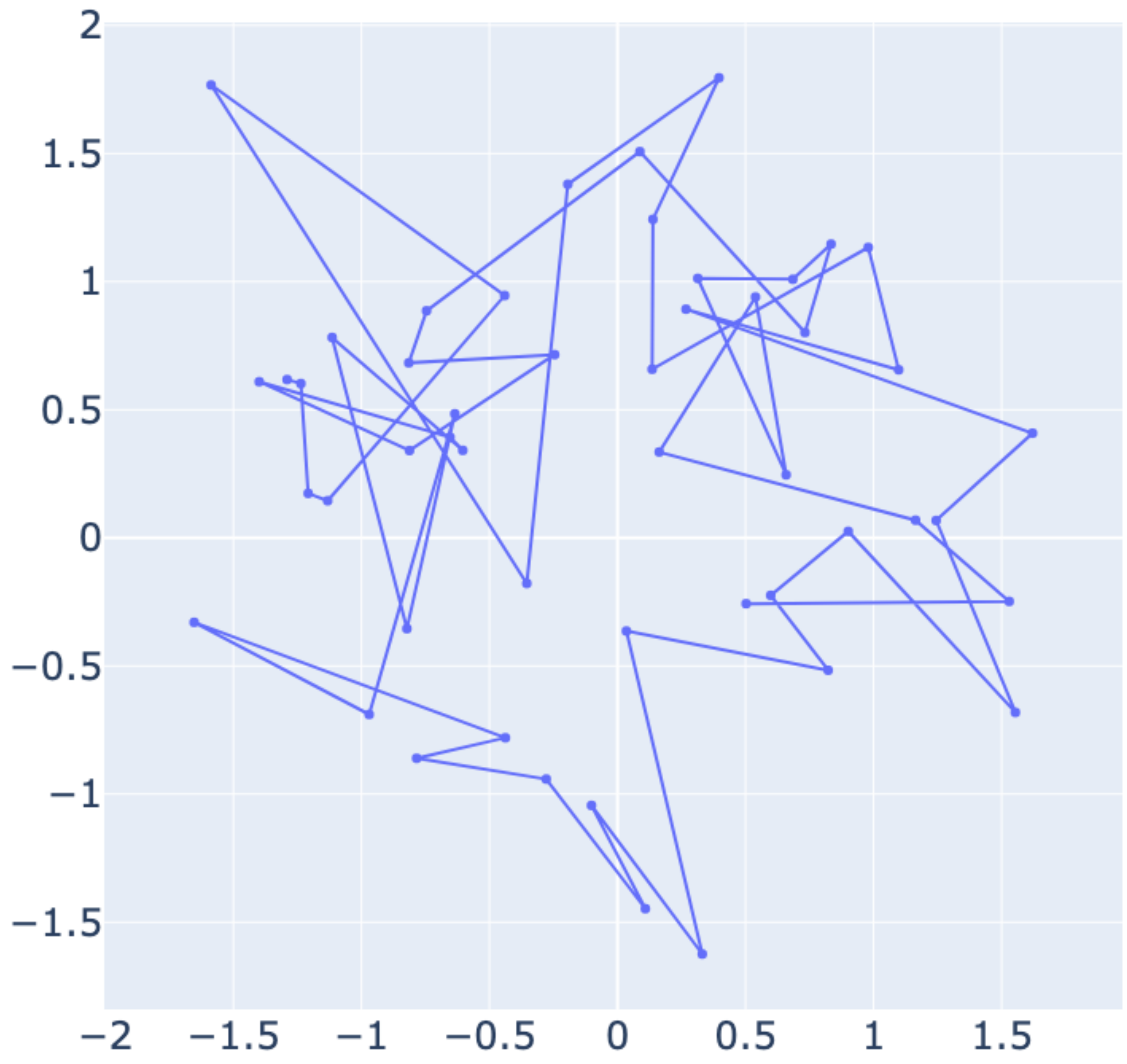
iid = independent identically distributed

Time series: observations

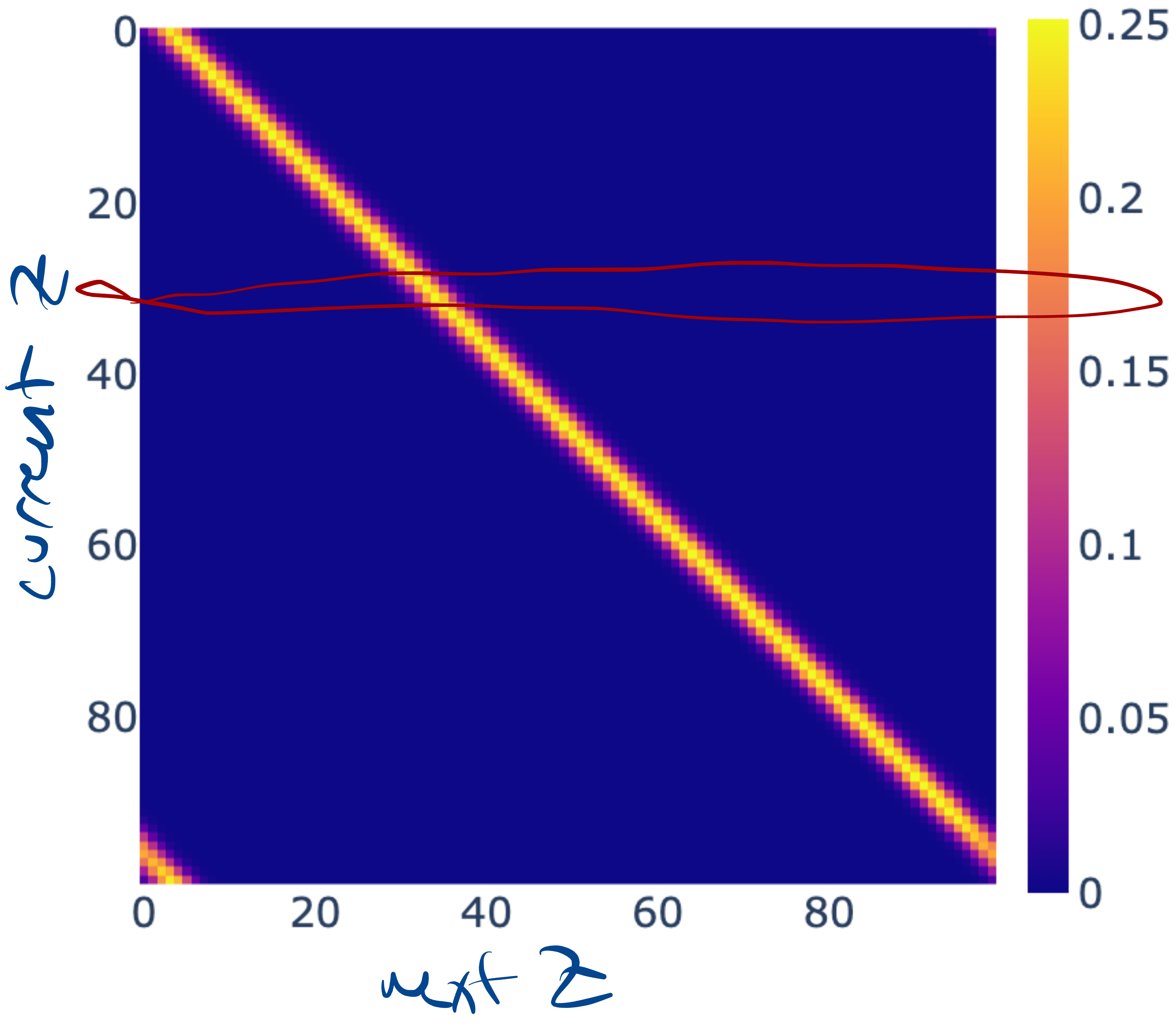
start: $P(z_1)$ — observations: $P(x_1, y_1 \mid z_1)$ — transitions = how I move: $P(z_2 \mid z_1)$

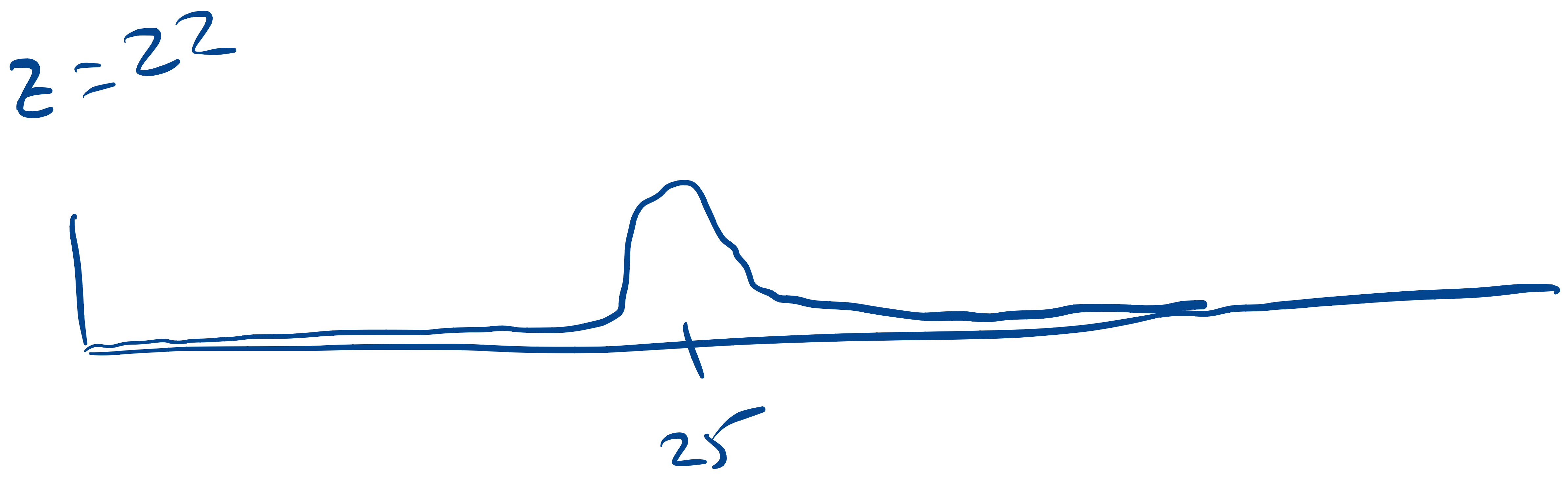
$$P(z_1)\,P(x_1, y_1 \mid z_1)\,P(z_2 \mid z_1)$$
$$P(x_2, y_2 \mid z_2)\,P(z_3 \mid z_2)$$

$\vdots \qquad \vdots$



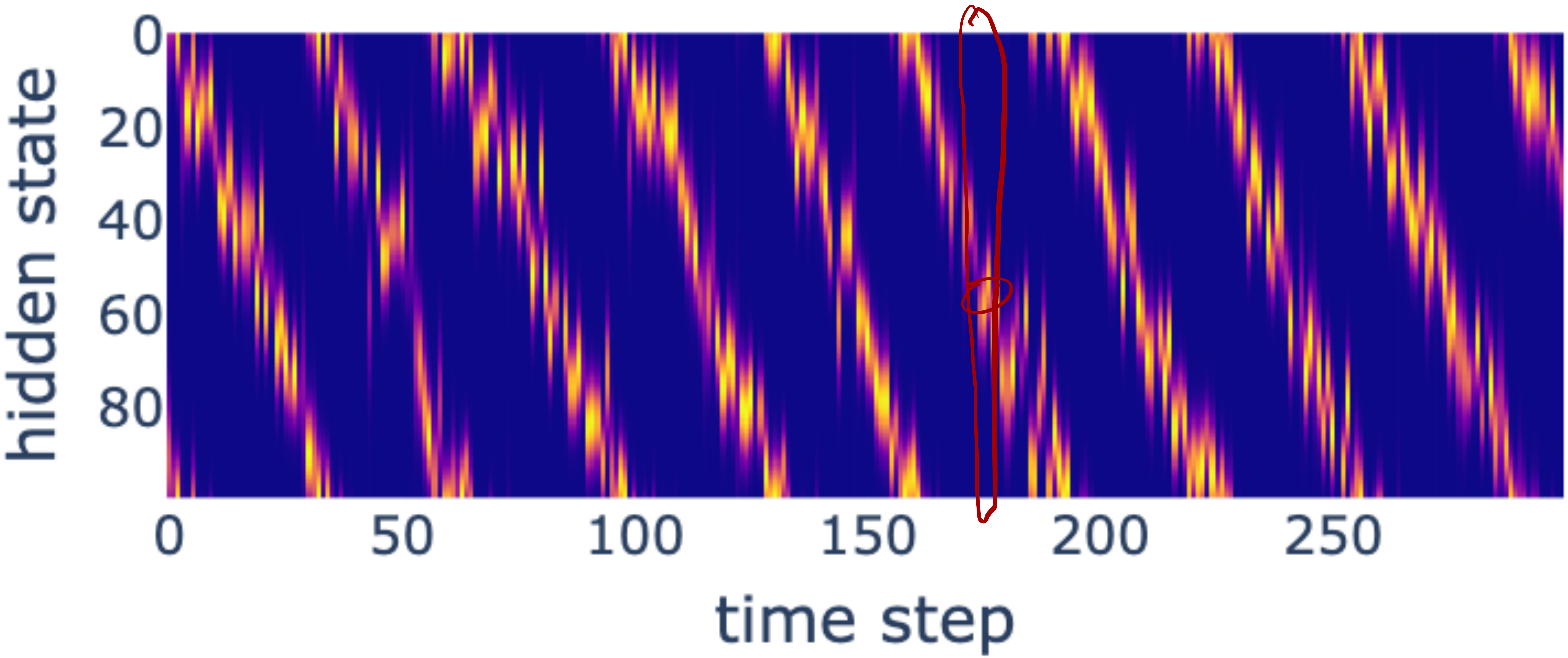
Model: transitions





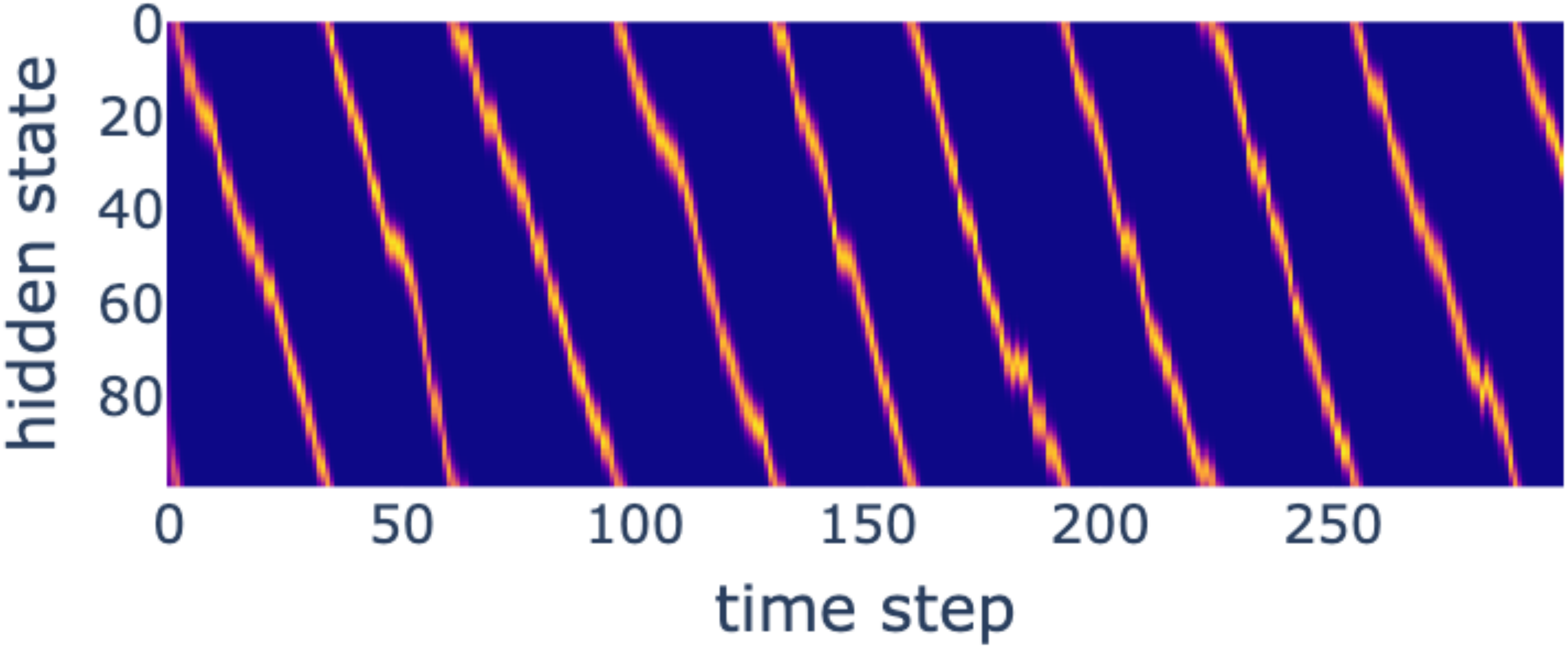
Model: observation likelihoods

P(observation | state) for given observations, all states



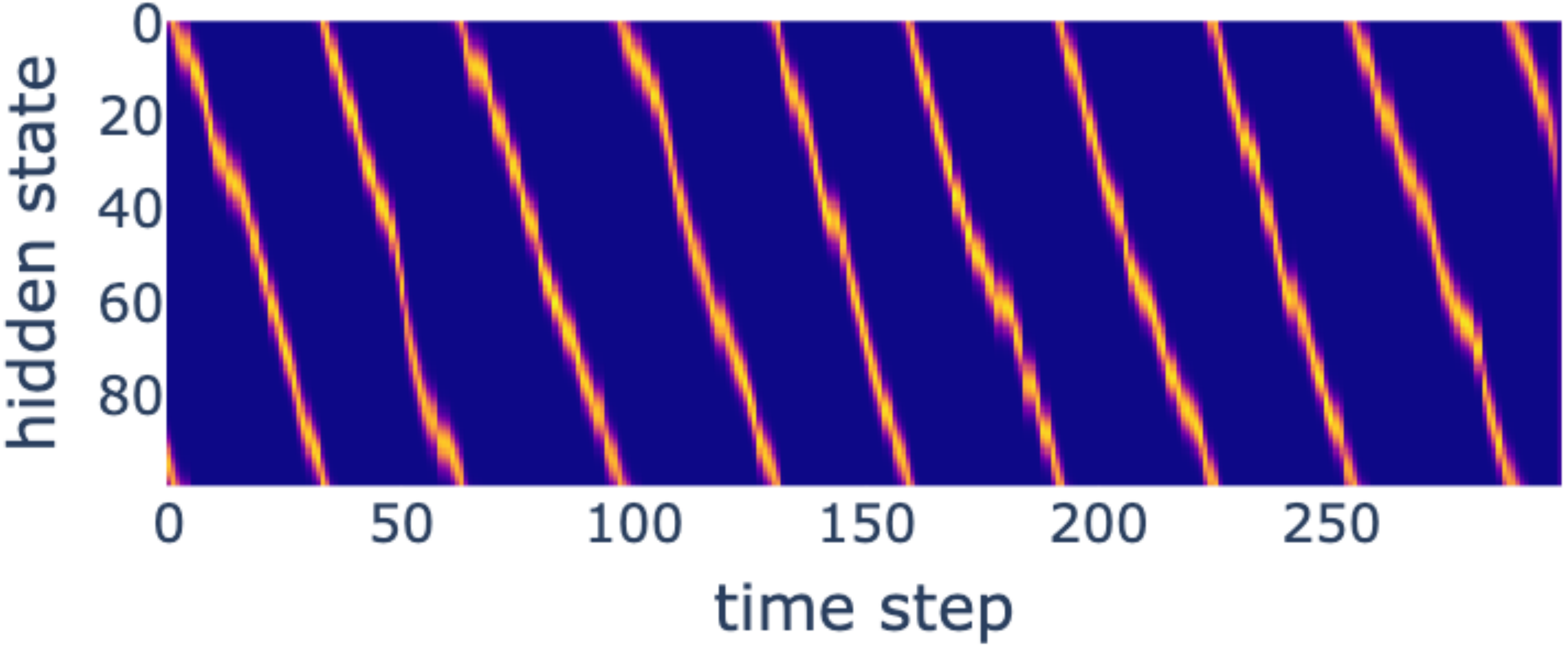
Inference: filtering distributions

P($state_t$ | all observations up to t)



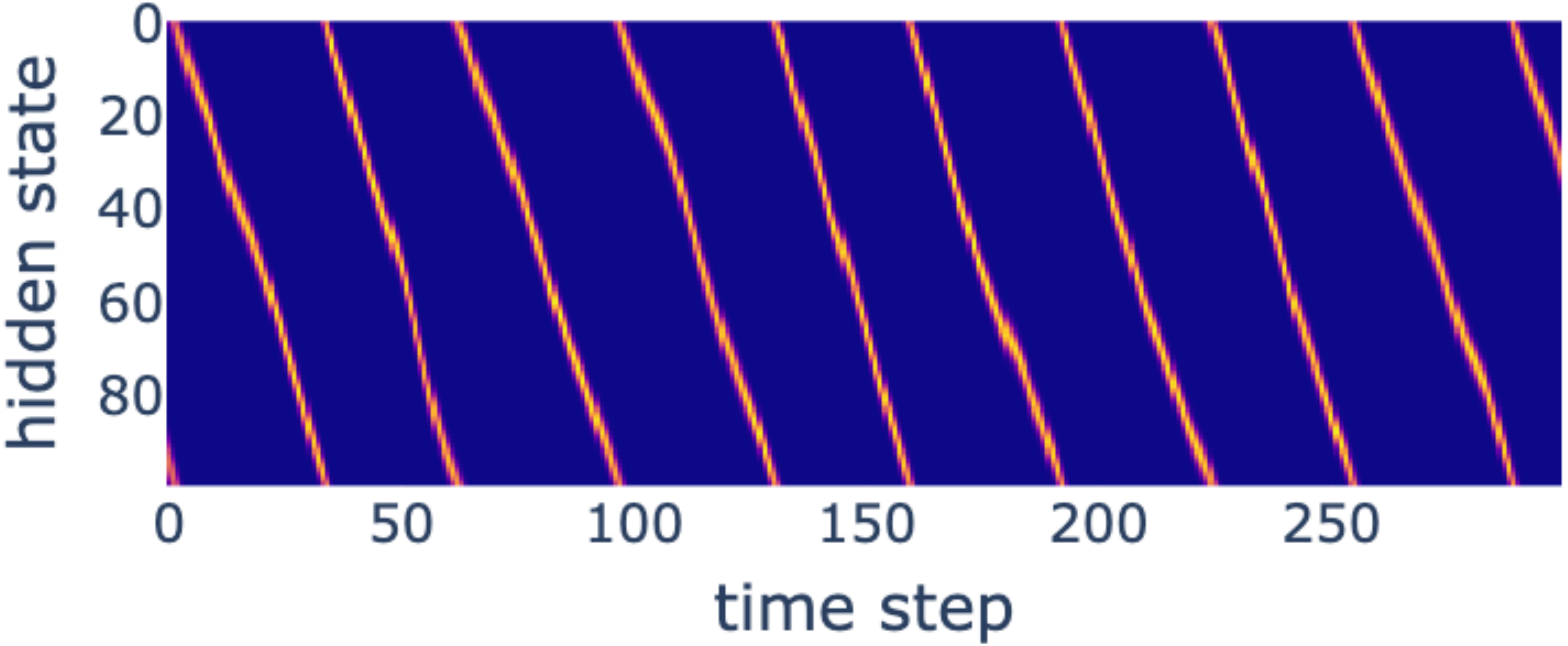
Backward filtering distributions

P(all observations after t | $state_t$), normalized



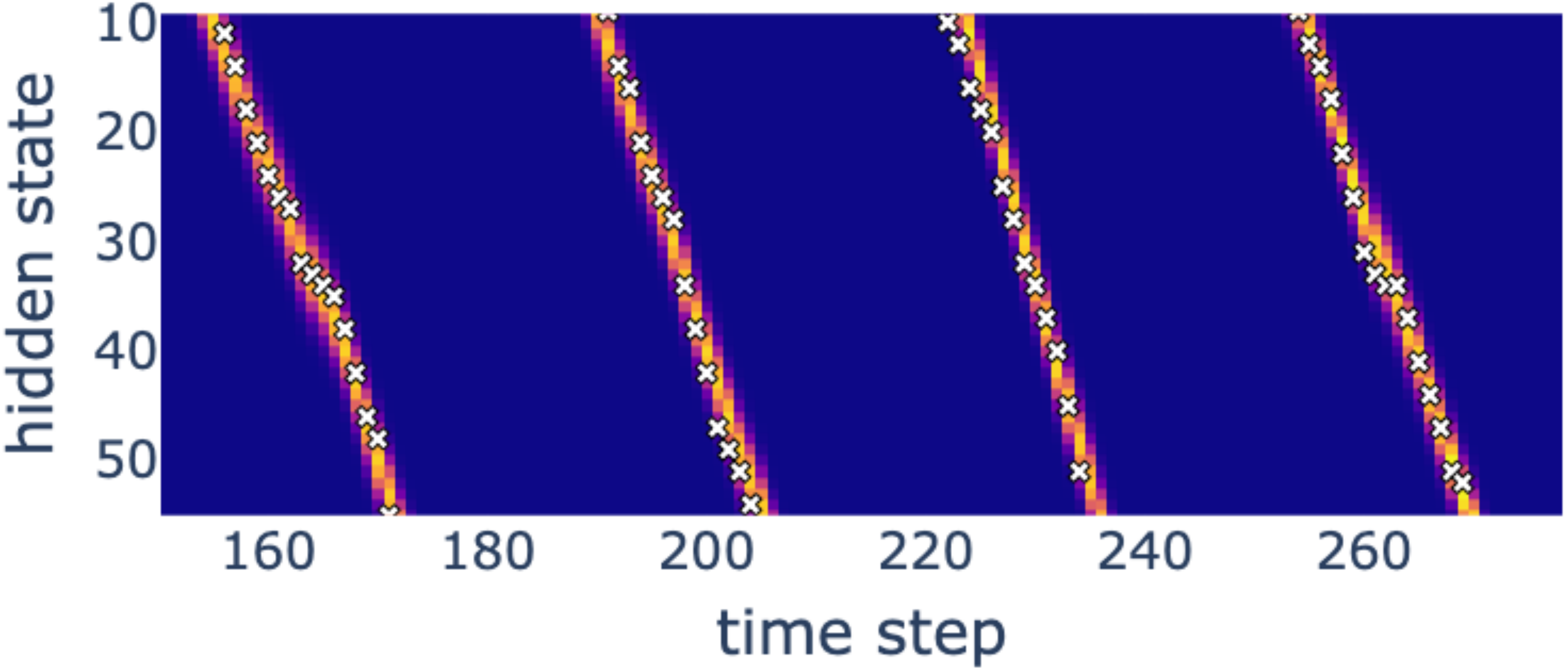
Marginal state posteriors

P($state_t$ | all observations)



Zoom of marginal state posteriors

with true hidden states marked



$$\min_{x,y} \quad \exp(x) + \exp(-x) + (x+y)^2 + 3x \quad \text{st} \quad x^2 + y^2 = 1$$

$$L = \exp(x) + \exp(-x) + (x+y)^2 + 3x + \alpha(x^2+y^2-1)$$

$$dL = \exp(x)\,dx - \exp(x)\,dx + 2(x+y)\,dx + 3dx + 2\alpha x\,dx = 0$$

$$dL = 2(x+y)\,dy + 2\alpha y\,dy = 0$$

$$dL = (x^2+y^2-1)\,d\alpha = 0$$

$$z = x^2 + y^2$$

$$x = \cos(t)$$

$$y = \sin(t)$$

$$dz = 2x\,dx + 2y\,dy$$

$$dx = -\sin(t)\,dt$$

$$dy = \cos(t)\,dt$$

$$dz = 2x(-\sin(t)\,dt) + 2y(\cos(t)\,dt)$$

$$= -2\cos(t)\sin t\,dt + 2\sin t\cos t\,dt$$

$$= 0$$

$$f \in \mathbb{R}^n \to \mathbb{R}^{k \times d}$$

$$f'(x) \in \underbrace{\mathbb{R}^n \to \mathbb{R}^{k \times d}}$$

$$f''(x) \in \underbrace{\mathbb{R}^n \to \mathbb{R}^n \to \mathbb{R}^{k \times d}}$$

$$df = \underbrace{f'(x)}\, dx$$

$$d^2 f = dx \cdot \underbrace{f''(x)}\, dx$$