Schrödinger Bridges & CLIP

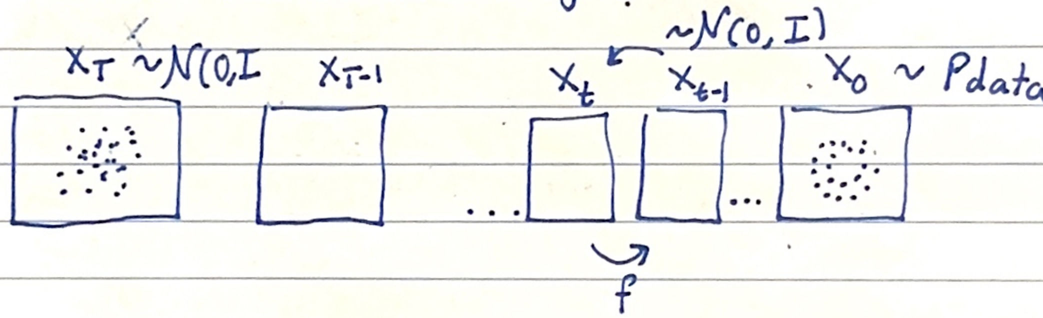
Plan

Review
Schrödinger Bridges
CLIP

Logistics

Games @ 5pm Tomorrow
Strict deadline (unless we talk)
Visit for Note Feedback

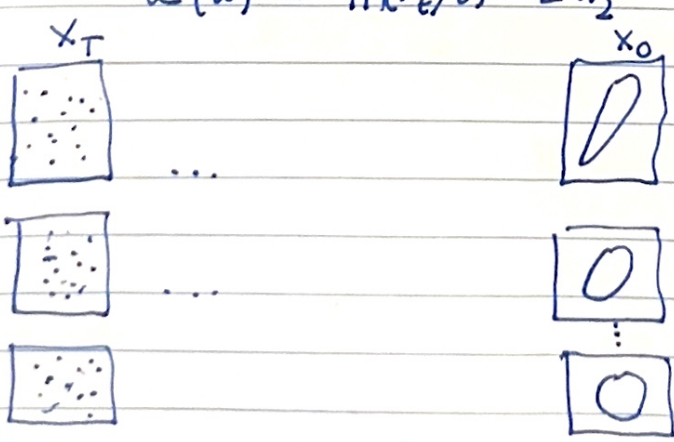
Motivation: Generate new images!



$$x_t = \sqrt{\alpha^t}\, x_0 + \sqrt{1-\alpha^t}\, z \quad \text{for} \quad z \sim \mathcal{N}(0,I)$$

Train:

Sample $z \sim \mathcal{N}(0,I)$
Sample $t \sim \mathrm{Unif}([0,T])$
$x_t = \sqrt{\alpha^t}\, x_0 + \sqrt{1-\alpha^t}\, z$
$\mathcal{L}(w) = \| f(x_t, t) - z \|_2^2$

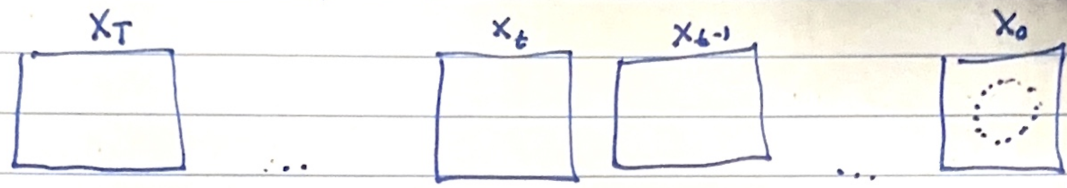


Eval:

helps if training x_T in is not $\mathcal{N}(0,I)$

$x_T \sim \mathcal{N}(0,I)$
for i in $\{1, \dots, T\}$:
$Z\_pred = f(x_T, T)$
$$x_0 = \frac{x_T - \sqrt{1-\alpha^T}\, Z\_pred}{\sqrt{\alpha^T}}$$
$Z \sim \mathcal{N}(0,I)$
$x_T = \sqrt{\alpha^T}\, x_0 + \sqrt{1-\alpha^t}\, Z$

Motivation: Simpler (but less efficient) diffusion



$x_t = x_{t-1} + \gamma z$ for $z \sim N(0, I)$

Train:

$x_t \sim P_{data}$

for t in $\{0, \dots, T-1\}$:

$x_{t+1} = x_t + \gamma z$

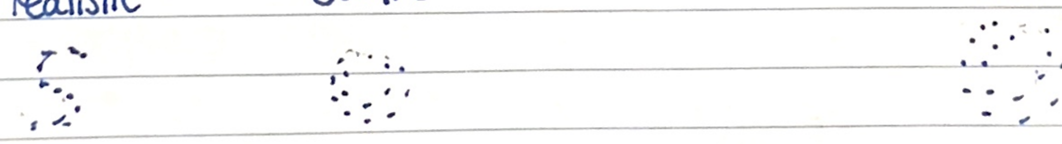
$\mathcal{L}(w) = \| f(x_{t+1}) - x_t \|_2^2$

Evaluation:

Schrödinger Bridges

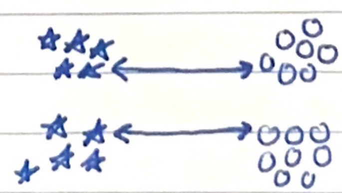
$P_{data'} \leftrightarrow P_{data}$ vs. $N(0, I) \leftrightarrow P_{data}$

e.g.,	blurry	clear	clear
	day	night	night
	realistic	comic	comic

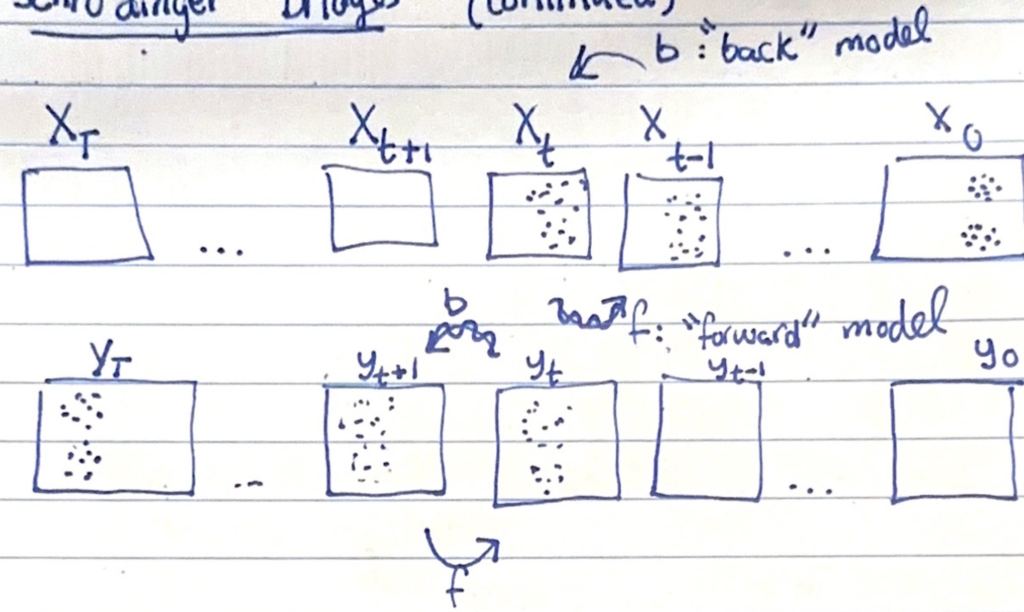


Motivation: Given $x_0 \sim P_{data}$, find closest $y_T \sim P_{data'}$

optimal transport :) e.g.,



Schrödinger Bridges (continued)



when 0, recover diffusion

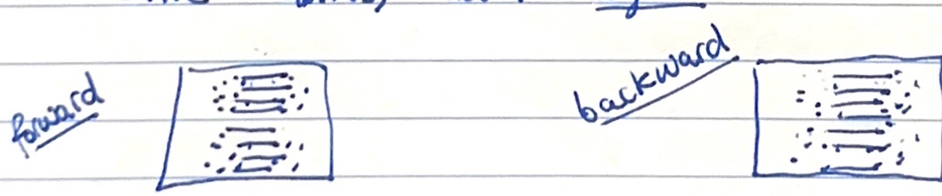
$$x_{t+1} = x_t + \gamma z + b(x_t^t)$$

$$\mathcal{L}_f = \|f(x_t^t) - x_{t-1}\|_2^2$$

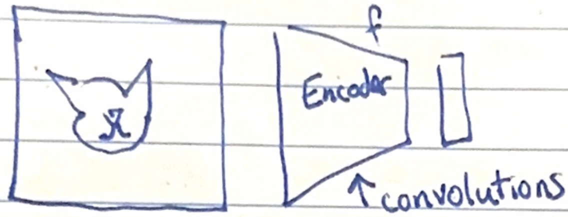
$$y_{t-1} = y_t + \gamma z + f(y_t^t)$$

$$\mathcal{L}_b = \|b(y_t^t) - y_{t+1}\|_2^2$$

Like GANs, train together



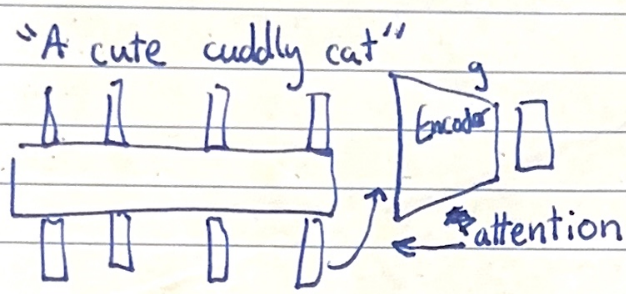
Contrastive Language Image Pretraining



Contrastive Learning

(image, caption) positive/negative

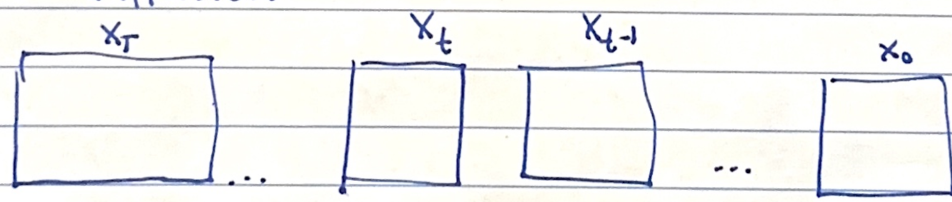
$|f(\text{image})^T f(\text{caption})|$ large/small



Immediate Application:

↳ Search for similar images/ captions

+ Diffusion:



$f(x_t, t, \boxed{\ })$ ← embedded caption