

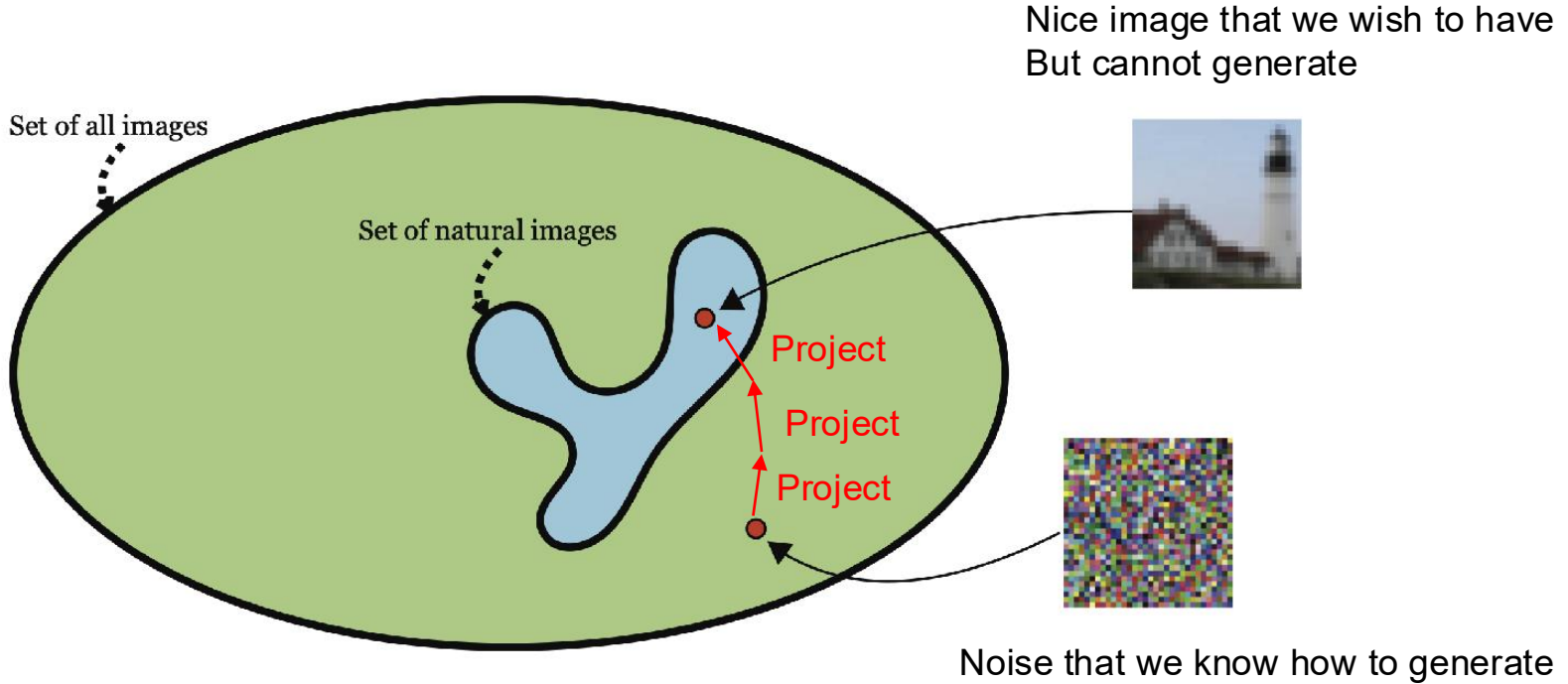
CS 180 Discussion 5

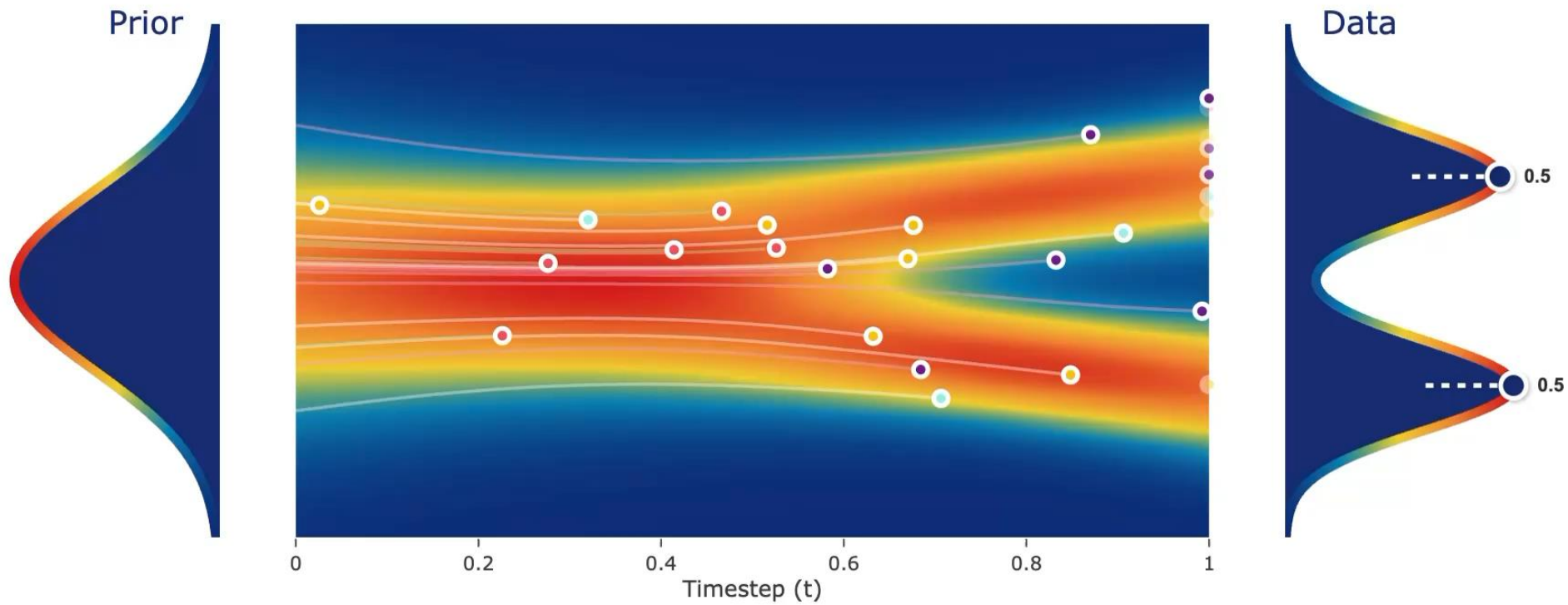
Flow Matching

Agenda

1. How does flow matching works?
2. (Maybe) Guidance

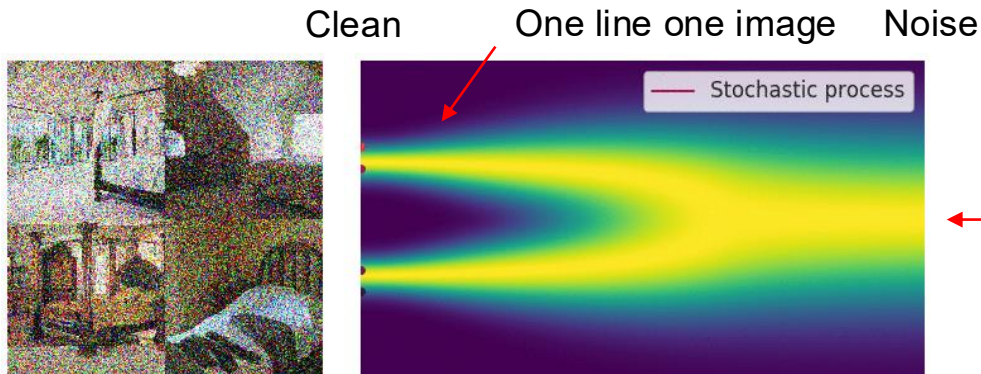
"Projection" to the natural image manifold



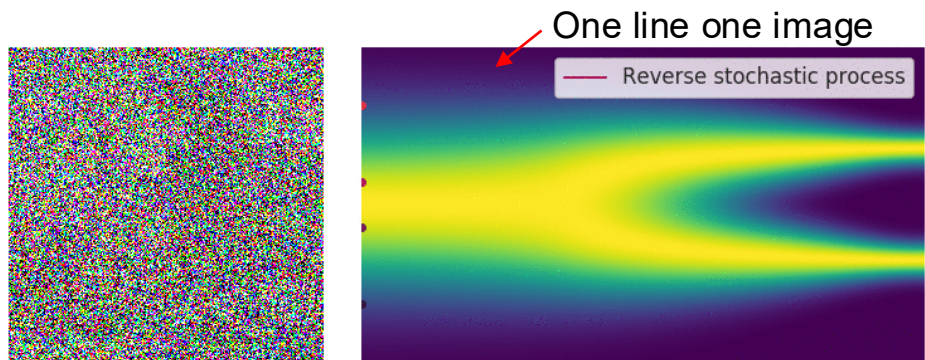


Diffusion process makes the “project” possible!

Forward
(adding noise)



Reverse
(denoising)



Noise

Clean

How to properly add noise and denoise?

(If it's that easy we wouldn't wait until 2020 to make it work)

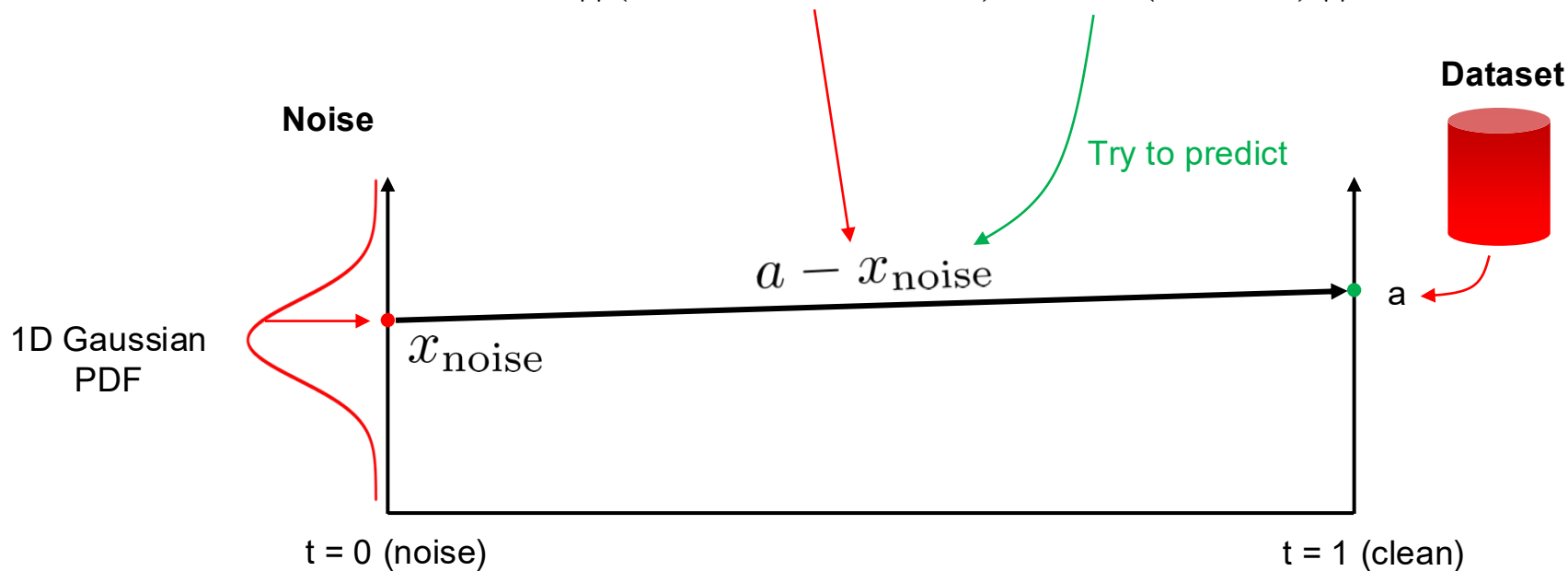
**To really understand what it learns,
Imagine a 1D world**

1D world with 2 datapoints



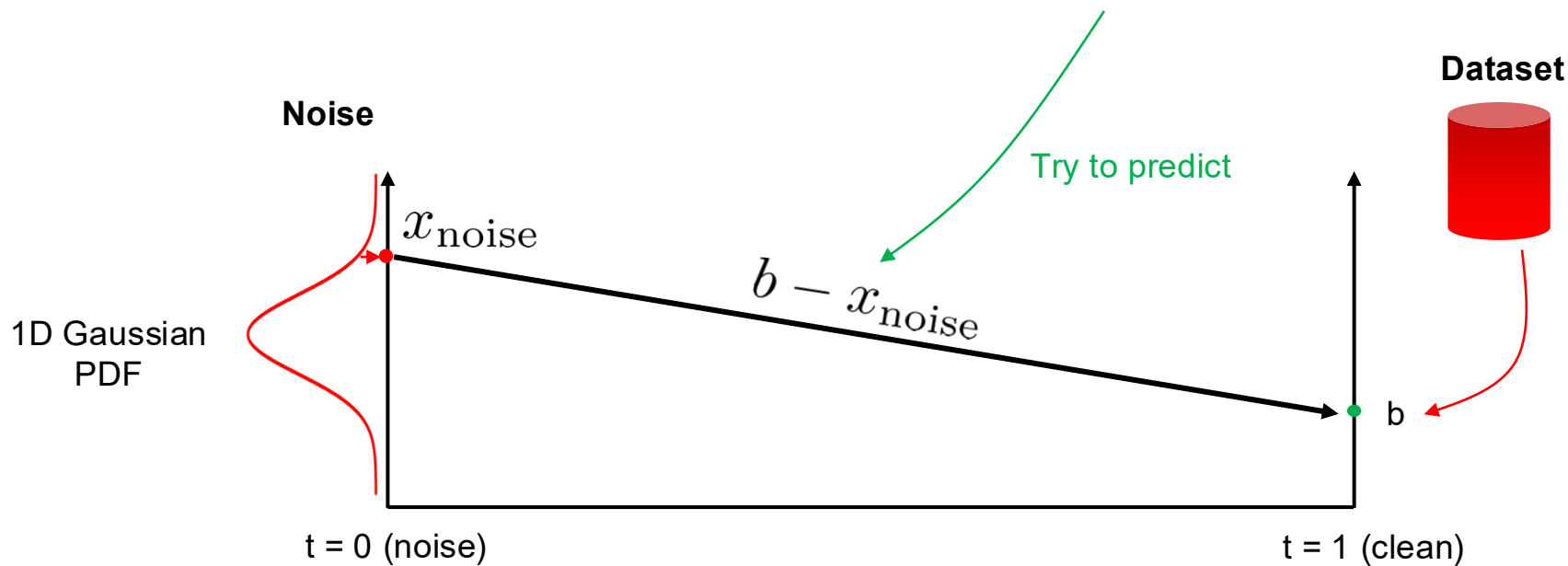
Sample noise, data, optimize

$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$



Sample noise, data, optimize

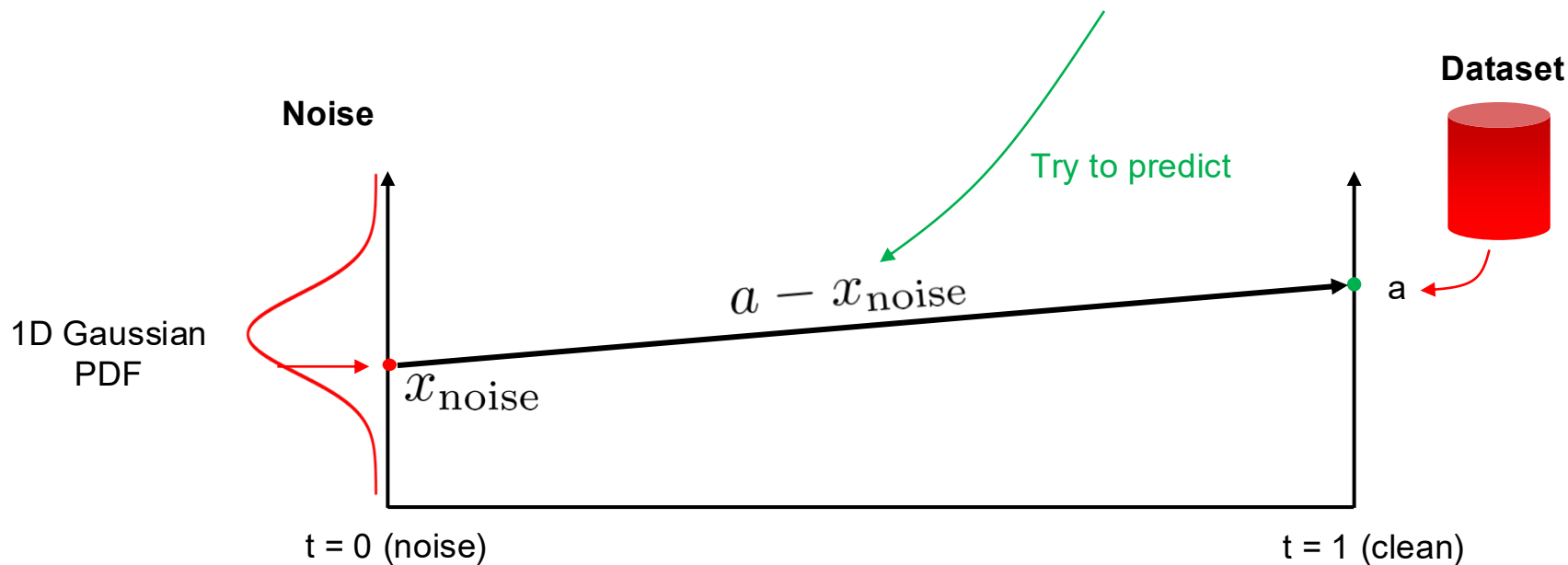
$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$



*Sample many times...

Sample noise, data, optimize

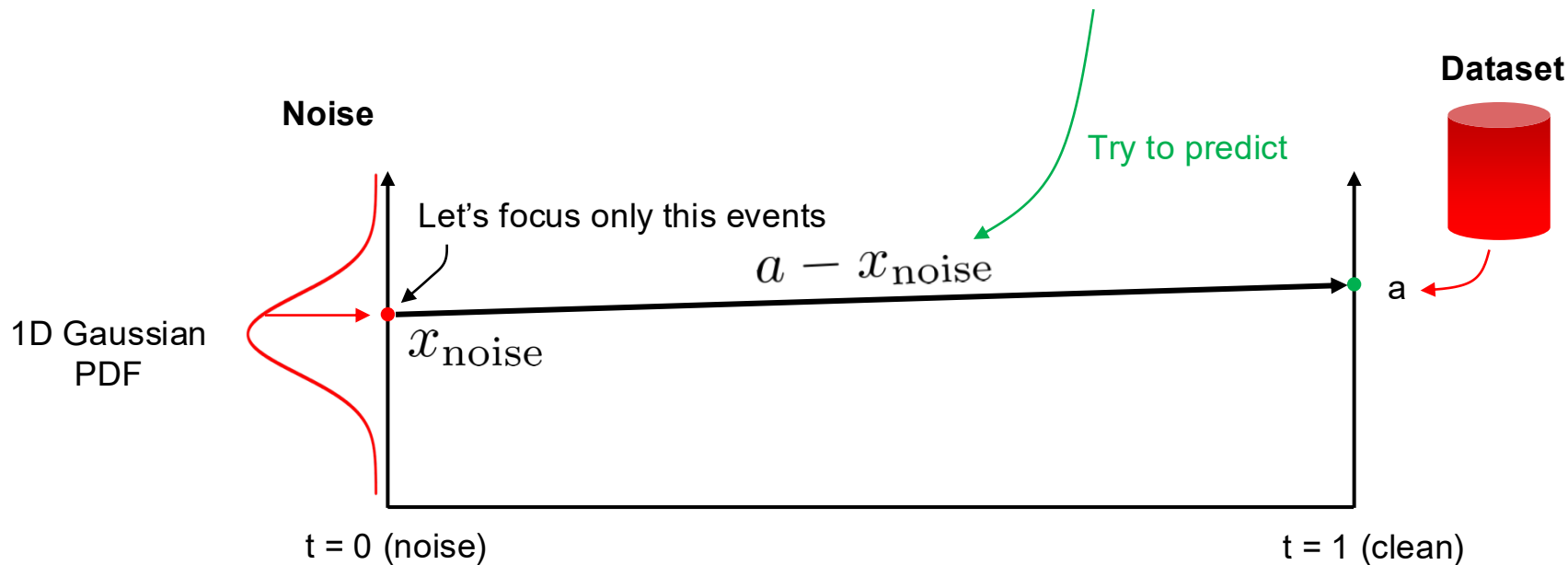
$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$



*Sample many times...

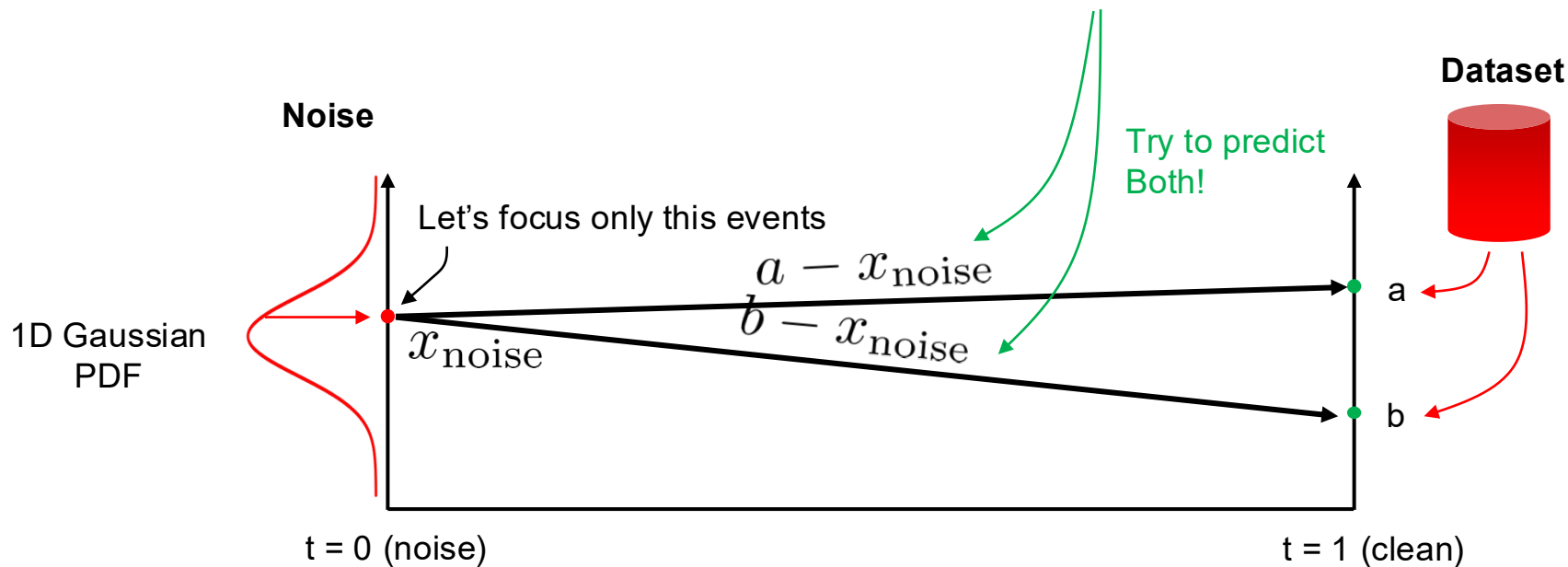
Let's focus on one event!

$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$



We have two datapoints!

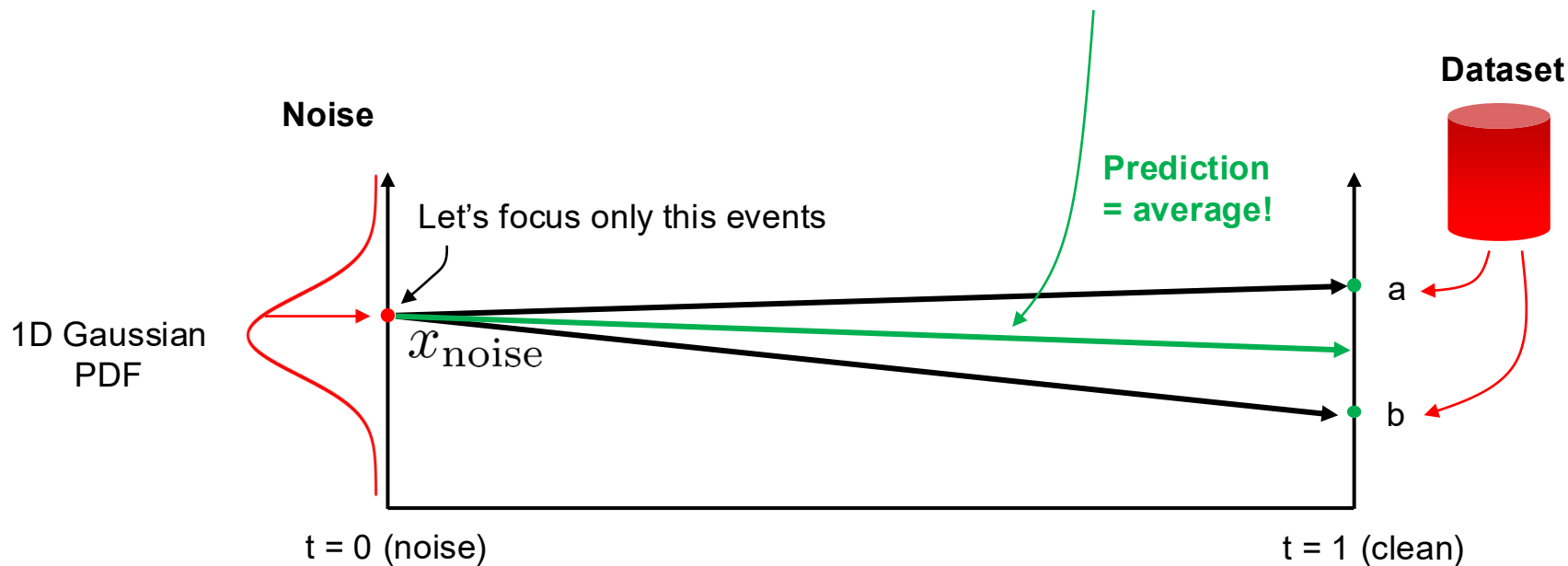
$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$



*Can neural net predict two values?

What will neural net predict? (Marginal flow)

$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_{\text{noise}})||^2$$

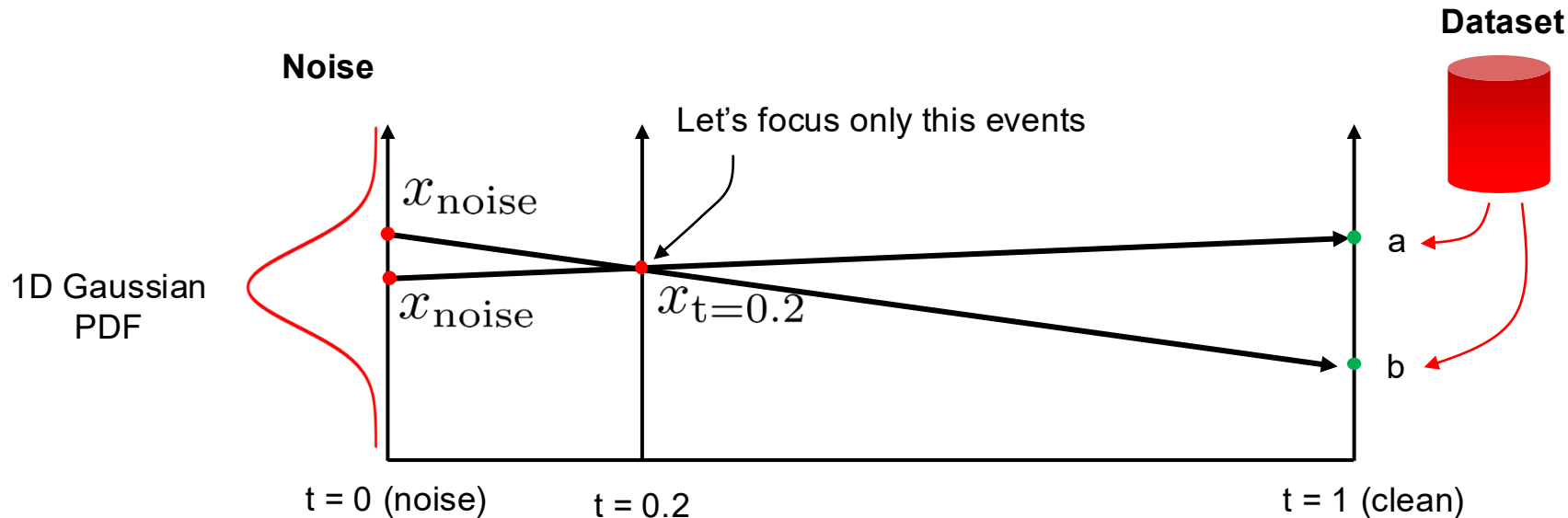


*At noise, the prediction is always the average!

Let's try at $t > 0$

What will neural net predict? (Marginal flow)

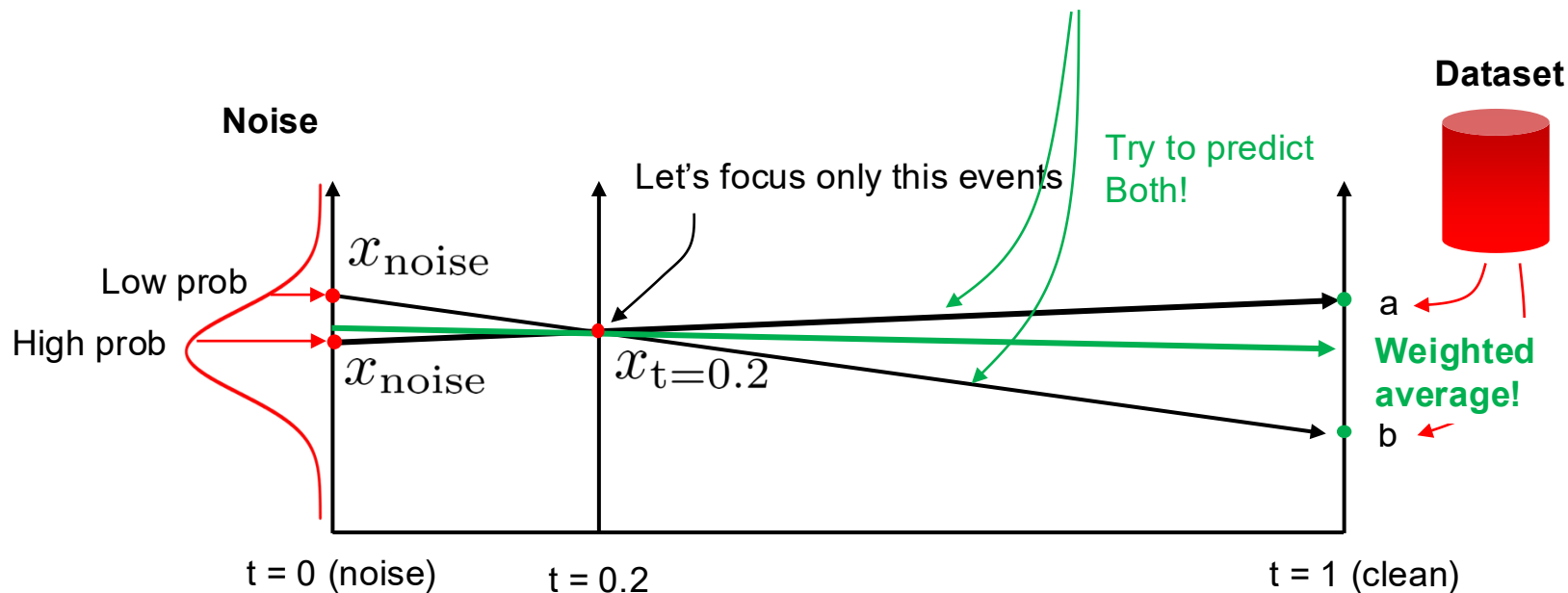
$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_t)||^2$$



Don't forget: $x_t = tx_{\text{clean}} + (1 - t)x_{\text{noise}}$

What will neural net predict? (Marginal flow)

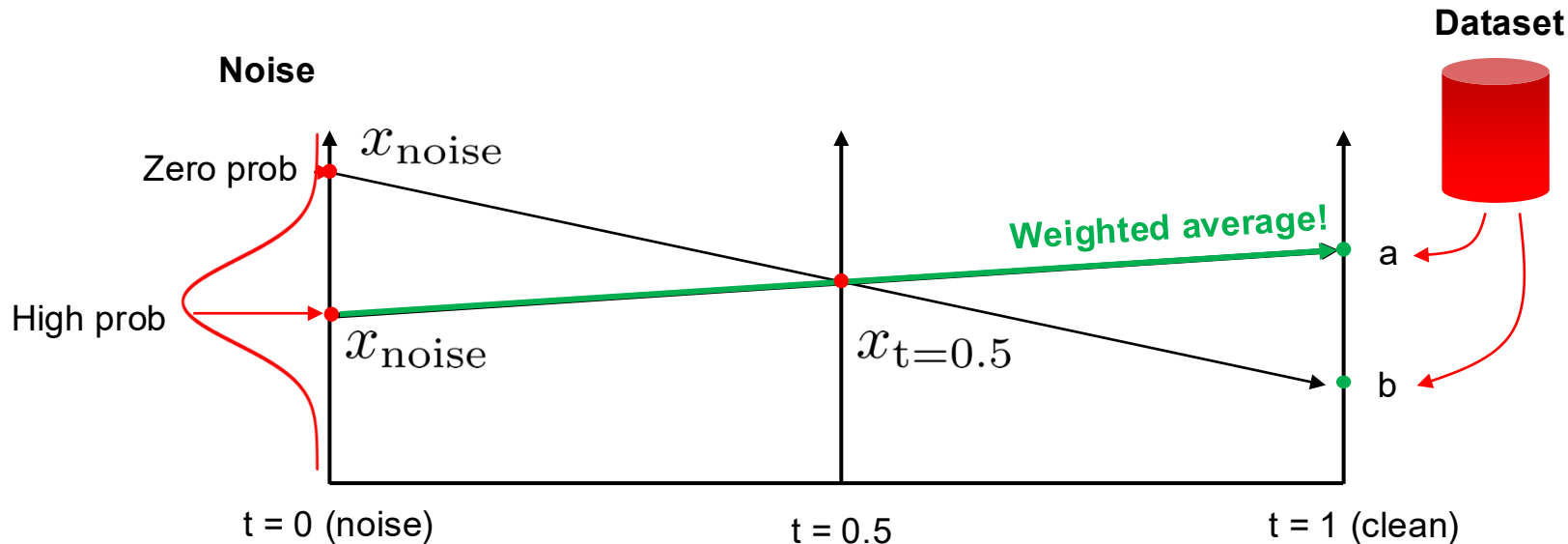
$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_t)||^2$$



Don't forget: $x_t = tx_{\text{clean}} + (1 - t)x_{\text{noise}}$

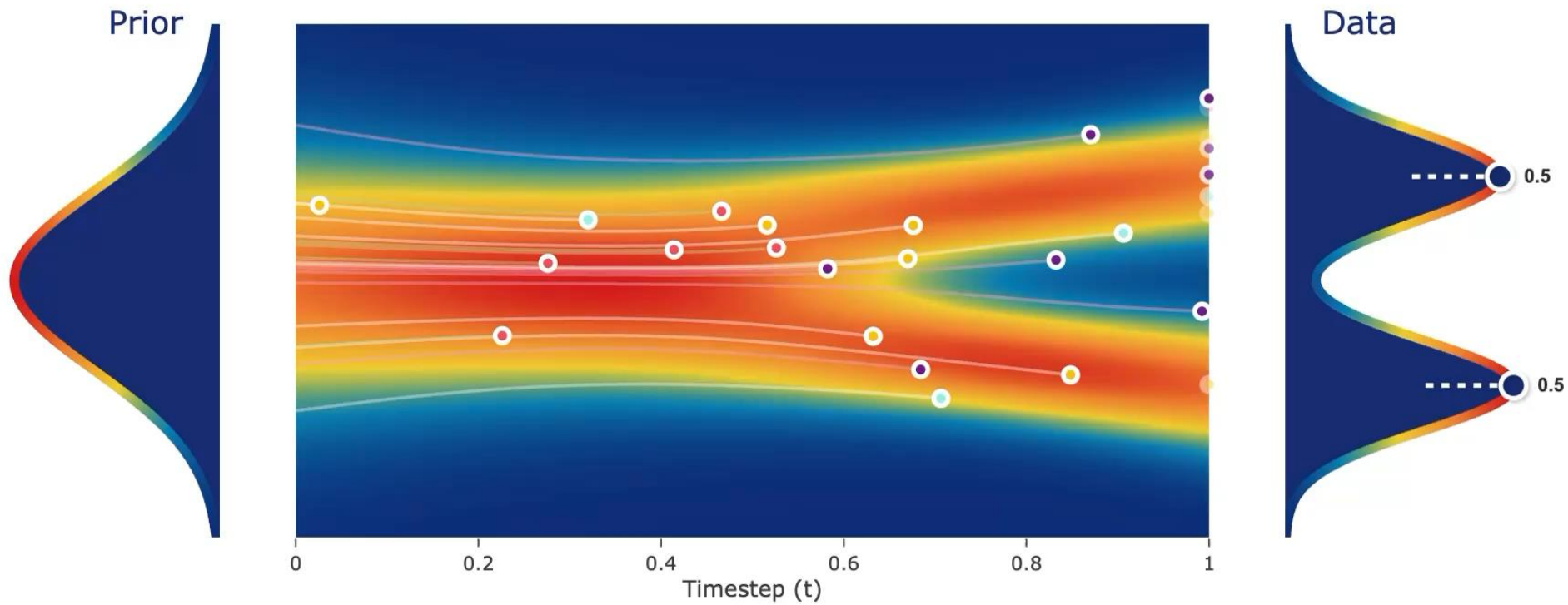
What will neural net predict? (Marginal flow)

$$L = ||(x_{\text{clean}} - x_{\text{noise}}) - u_{\theta}(x_t)||^2$$



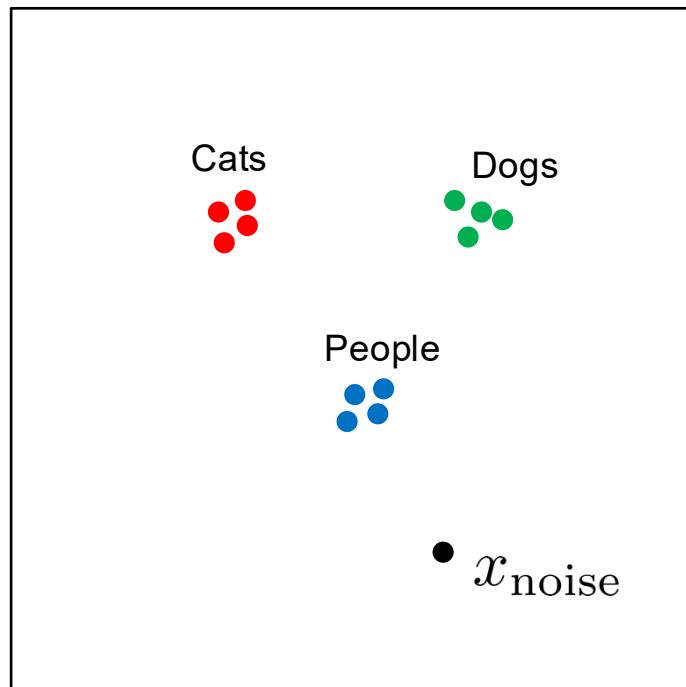
Don't forget: $x_t = tx_{\text{clean}} + (1 - t)x_{\text{noise}}$

*As t moves closer to clean, it is biased more to a real data point.

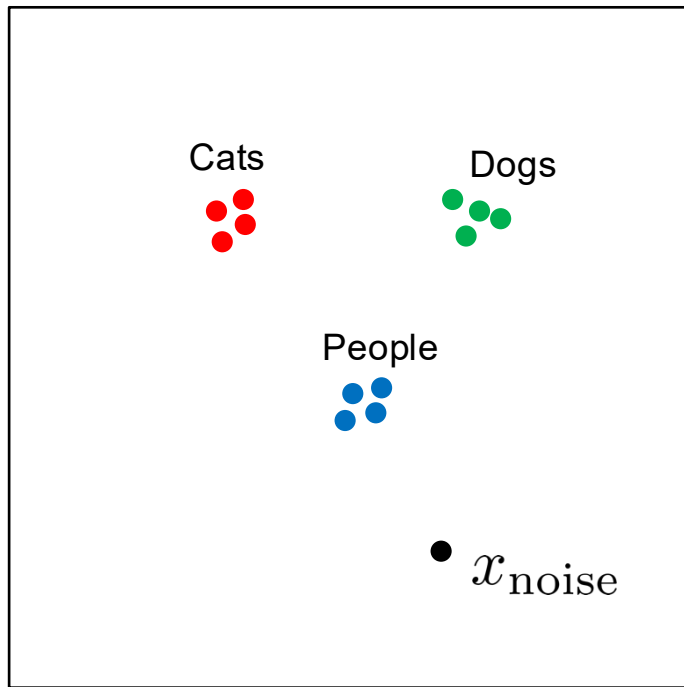


Classifier-free guidance (CFG)

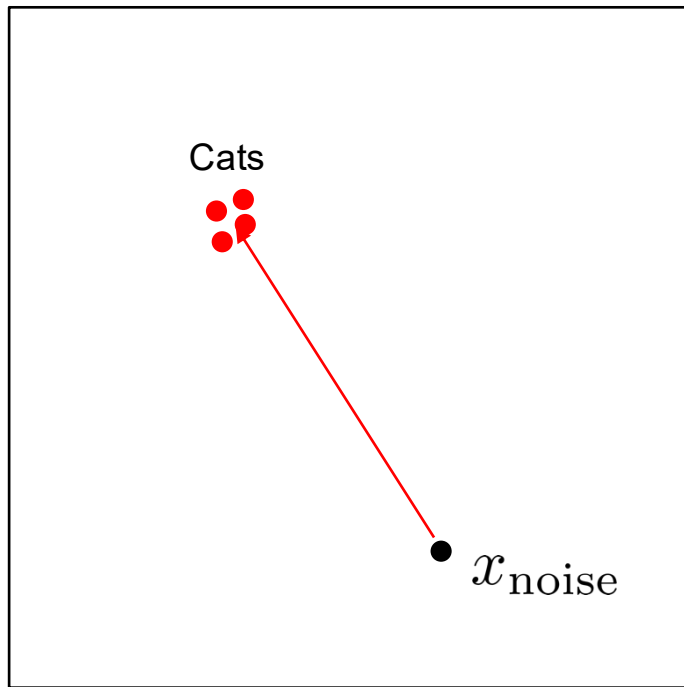
Now, imagine 2D world with 3 types of points



I want a "Cats" model, what's the marginal flow?

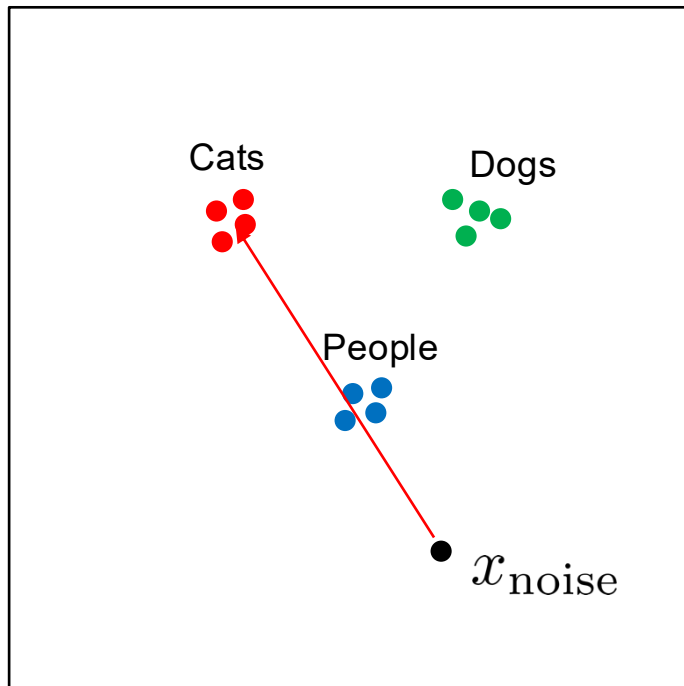


I want a "Cats" model, what's the marginal flow?



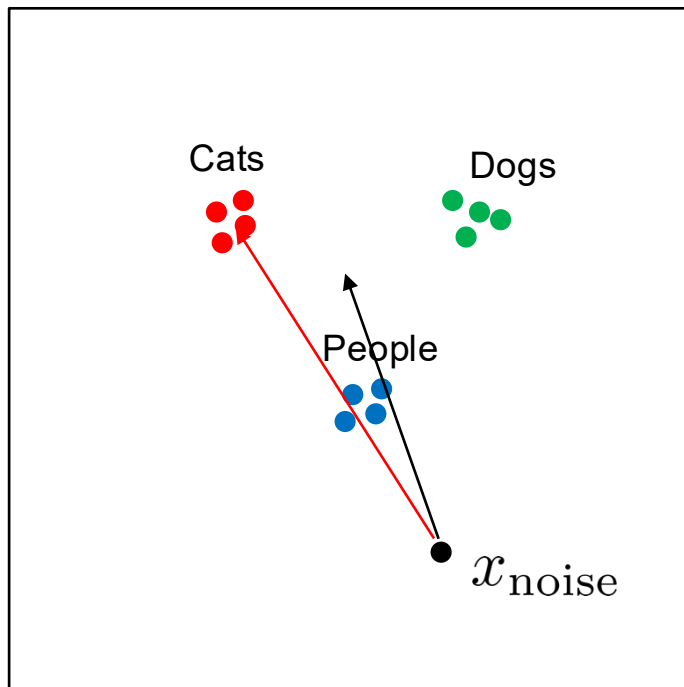
*At noise, point to the average!

What's unconditional marginal flow?



*At noise, point to the average!

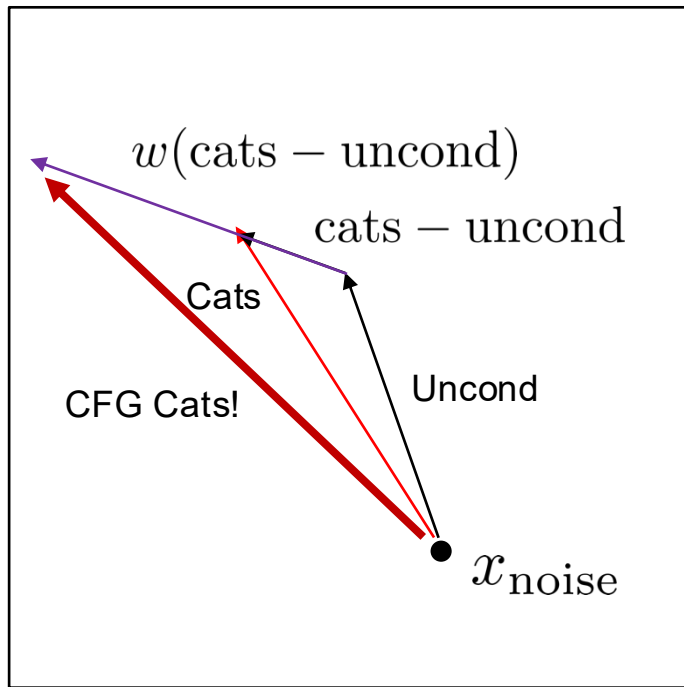
What's unconditional marginal flow?



*At noise, point to the average!

CFG = Exaggerate the "class" more

$$\text{CFG} = \text{uncond} + w(\text{cats} - \text{uncond})$$



Let's do worksheets!

*Why? People like catty-images not cat-ish images