

How to model $p(x)$

- $\log \hat{p}(x) = \text{CNN}(x)$ Infeasible
- Use the chain rule Cool, but slow to generate
- $p(x_1, \dots, x_d) = p(x_1) \dots p(x_d)$ Too restrictive
- Mixture of several Gaussians (GMM) Still too restrictive
- Mixture of infinitely many Gaussians

$$p(x) = \int p(x \mid t) p(t) dt$$