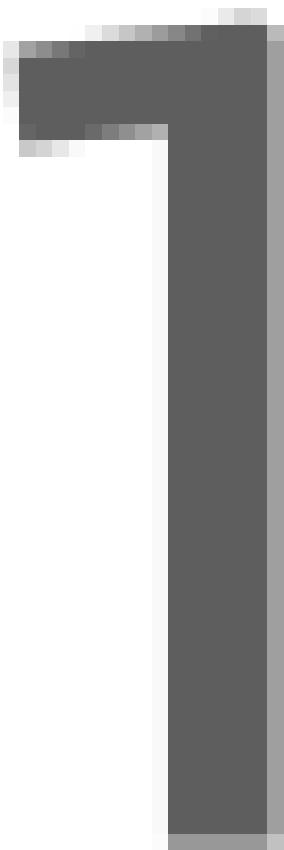
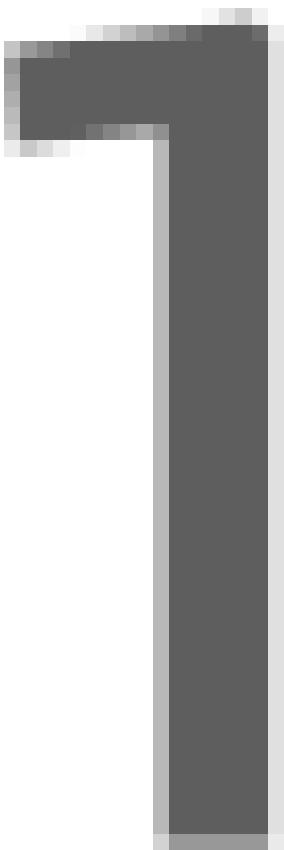


Siddhartha Mishra (MIT/Alfi) Summer School





The course of dimensionality

Where is most of the probability mass concentrated in high dimensions?

*Dealing with
high-dim data*



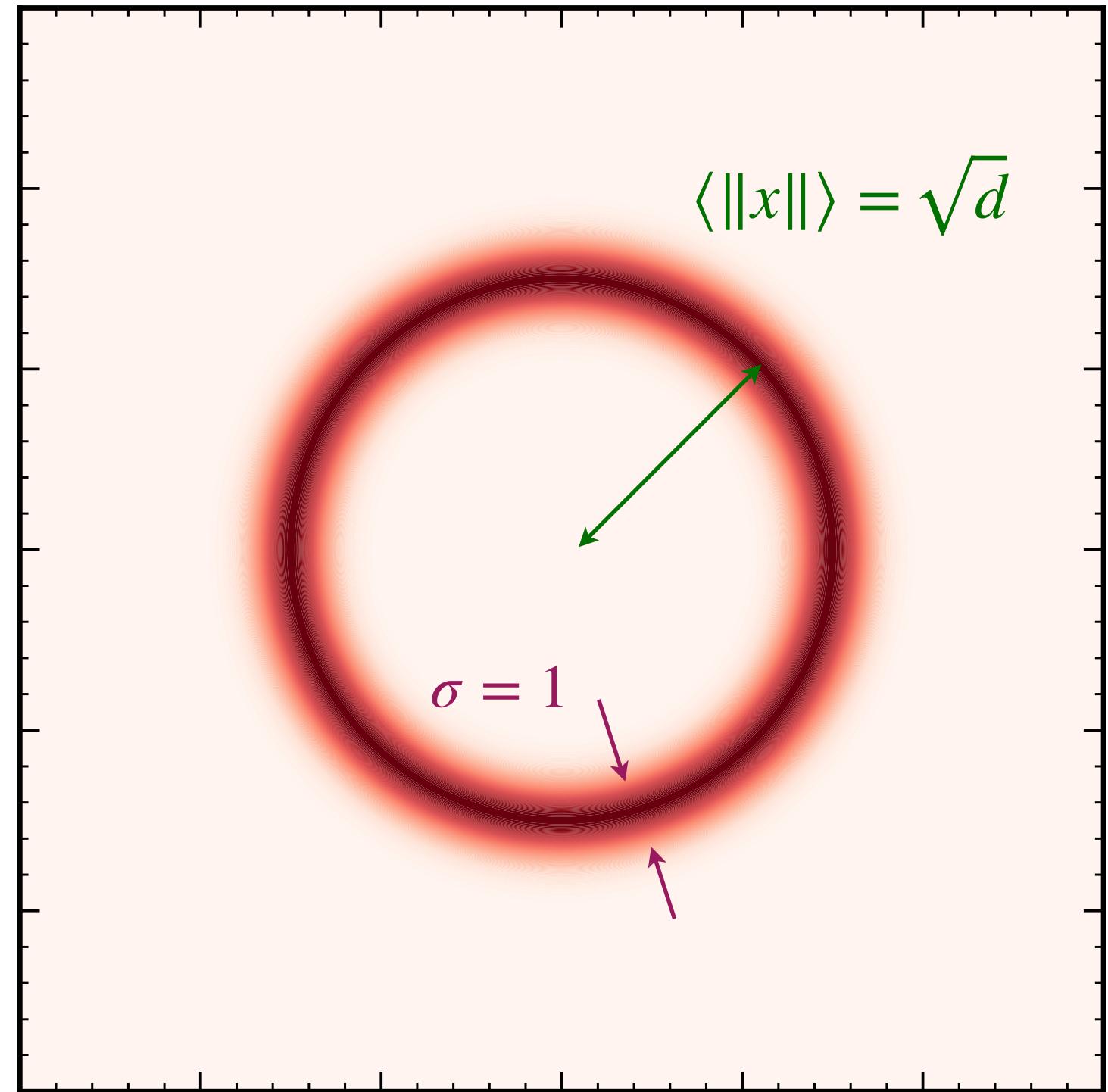
*Curse of
dimensionality*



Learning in high-dimensional spaces is challenging.

$$x \sim \mathcal{N}(0, \mathbb{I}_d)$$

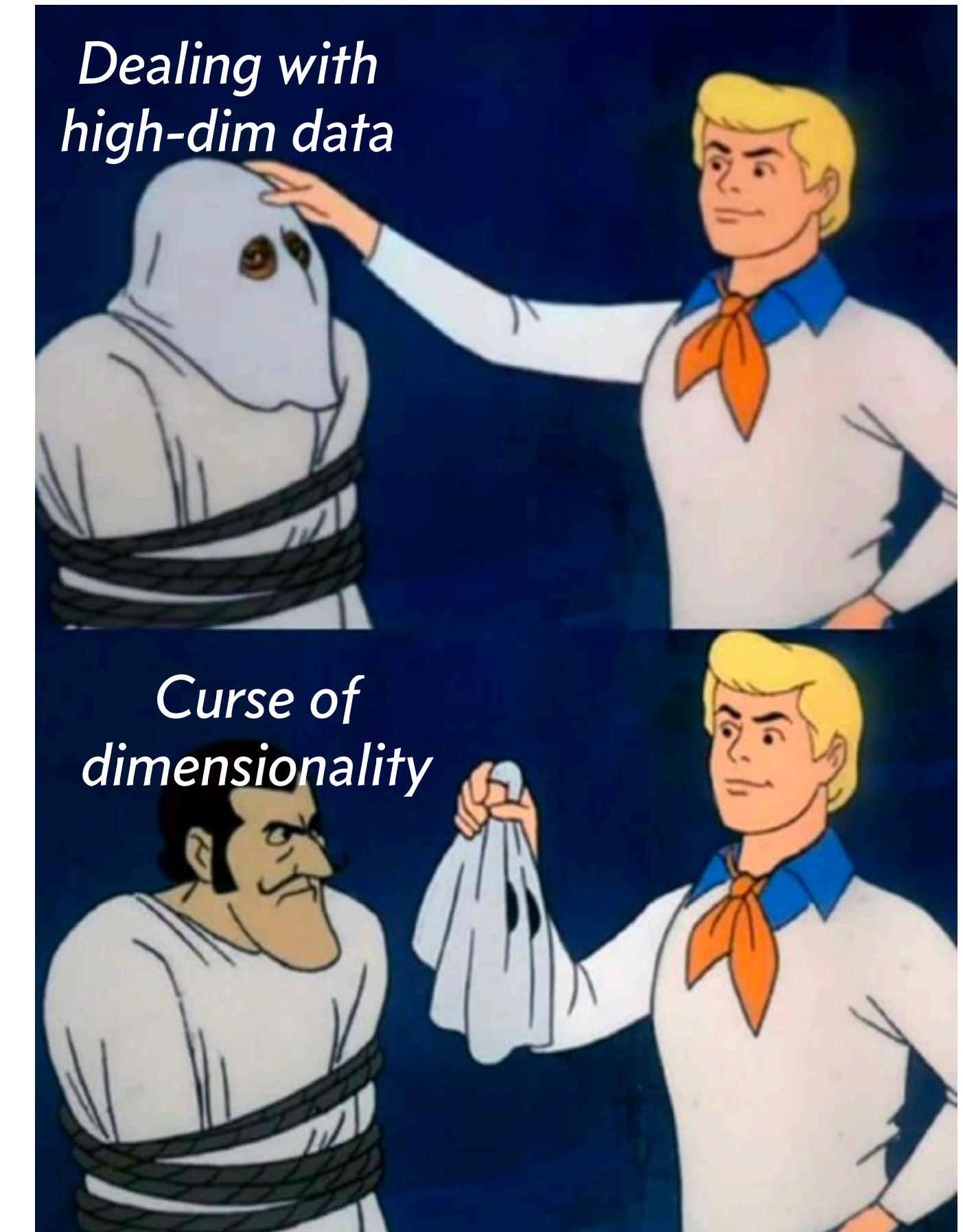
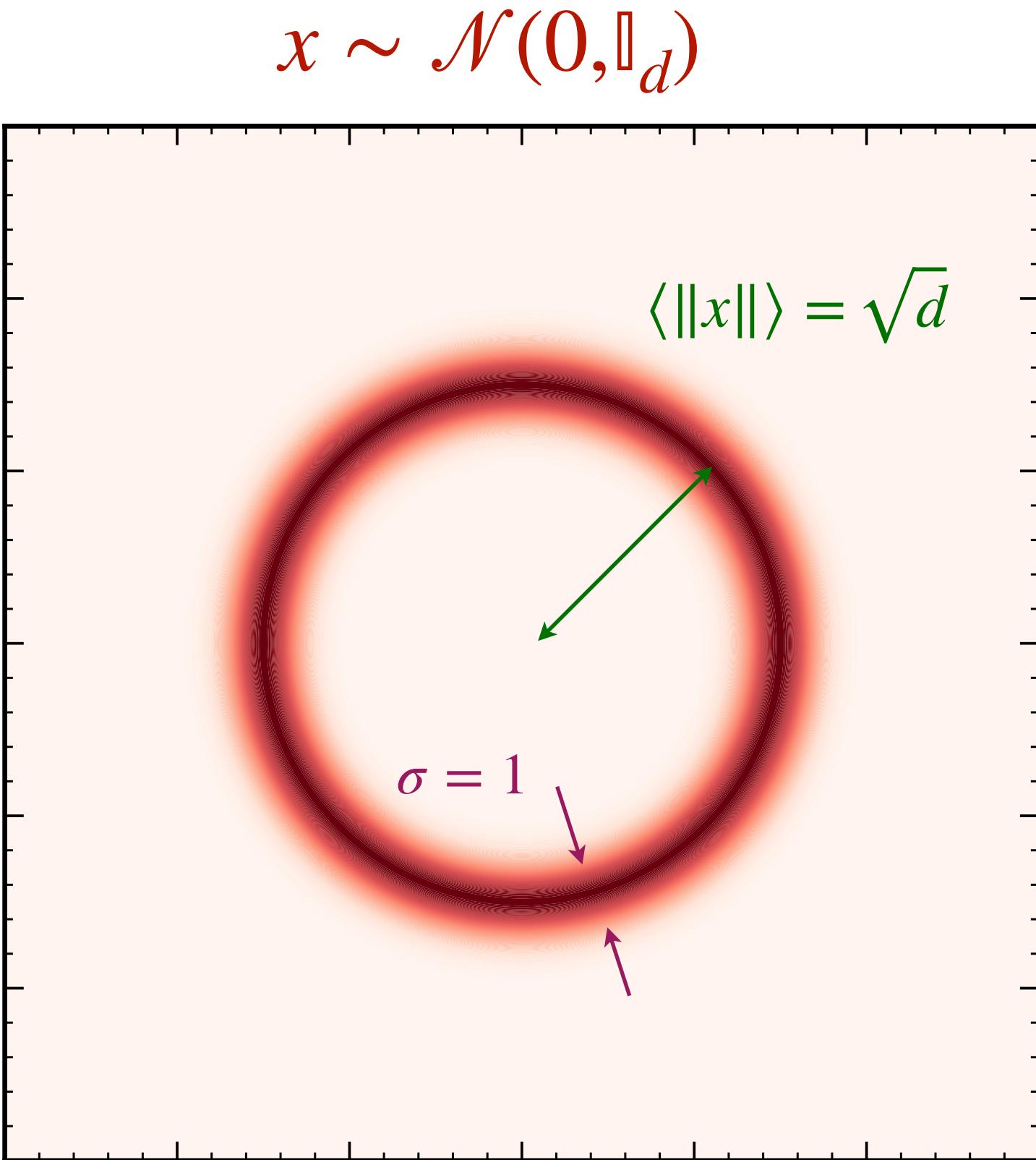
- In high dimensions, most of the probability density of a Gaussian distribution lies in a thin shell at distance \sqrt{d} from the center
- Vanishingly small fraction of distribution support is actually occupied.



The curse of dimensionality

Where is most of the probability mass concentrated in high dimensions?

- In high dimensions, most of the probability density of a Gaussian distribution lies in a thin shell at distance \sqrt{d} from the center
- Vanishingly small fraction of distribution support is actually occupied.



Learning high-dimensional distributions is challenging!

Typicality and likelihood of samples

Which of these samples have a higher likelihood under $\mathcal{L} = \mathcal{N}(0, \mathbb{I}_d)$?

