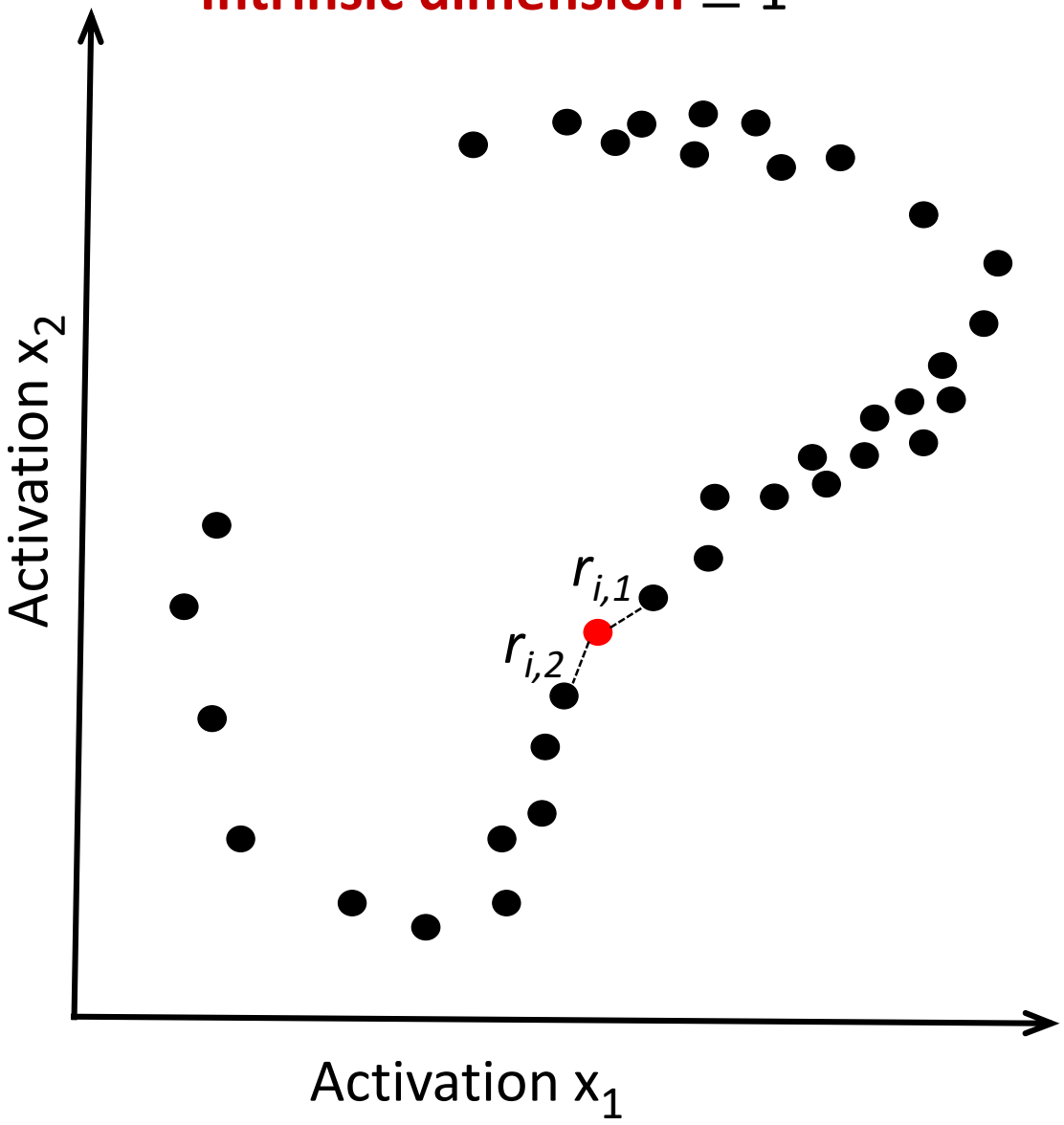
D = # of nodes in the layer = 2

Intrinsic dimension ≅ 1



1) For each data point *i* compute the distance to its first and second neighbour ($r_{i,1}$ and $r_{i,2}$)

2) For each *i* compute $\mu_i = \frac{r_{i,2}}{r_{i,1}}$

The probability distribution of μ is

$$P(\mu) = \frac{d}{\mu^{1+d}}$$

where d is the ID, independently on the local density of points.
3) Infer d from the empirical probability distribution of all the μ_i.

4) Repeat the calculation selecting a fraction of points at random. This gives the ID as a function of the scale.