

# Deficiency of PLSA

- Not a generative model
  - Can't compute probability of a new document (why do we care about this?)
  - Heuristic workaround is possible, though
- Many parameters → high complexity of models
  - Many local maxima
  - Prone to overfitting
- Overfitting is not necessarily a problem for text mining (only interested in fitting the “training” documents)