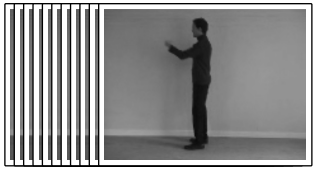
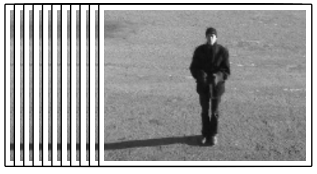
How to Supervise Topic Models

C. Zhang, and H. Kjellström

CVAP/CAS,KTH Royal Institute of Technology

About Classification


Boxing


Hand-clapping


Hand-waving

Action Classification


Headphone


Water Lilly


Lamp

Object Classification

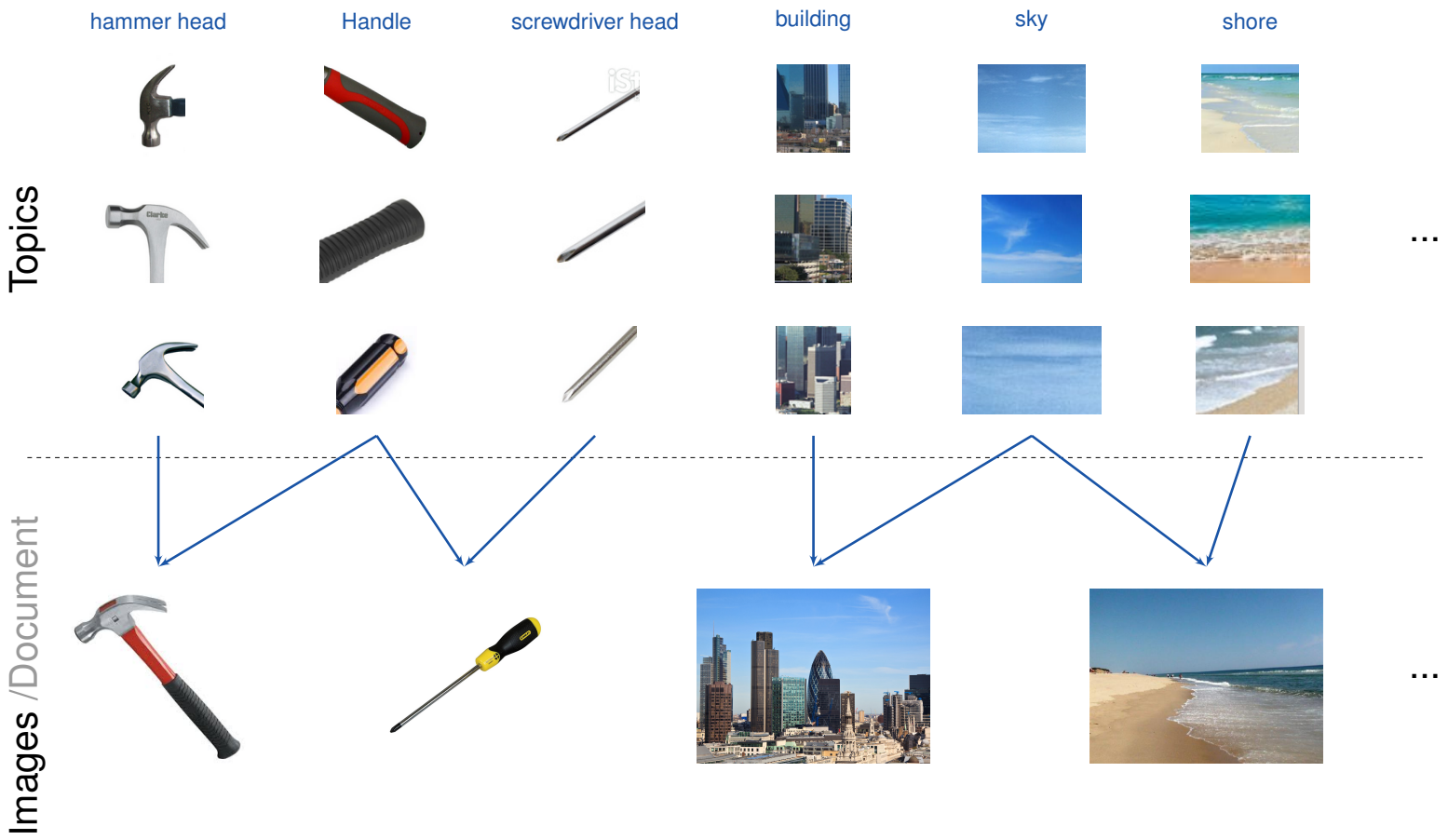

Mountain


Forest


Coast

Scene Classification

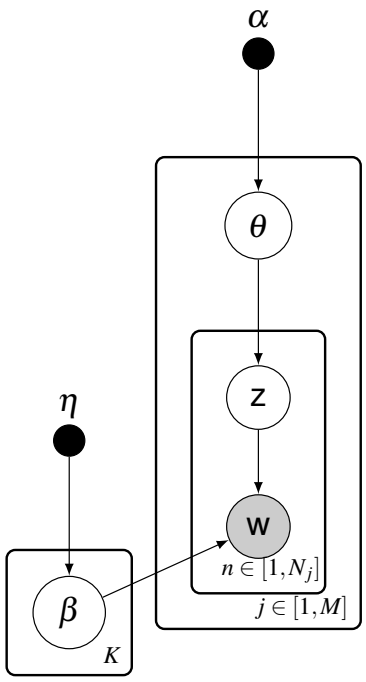
Introduction



How to supervise topic models?

1. Topic Models for classification
    a) Popular supervised topic models
    b) Novel variations of supervised topic models
2. Experiment evaluation
3. Future Work

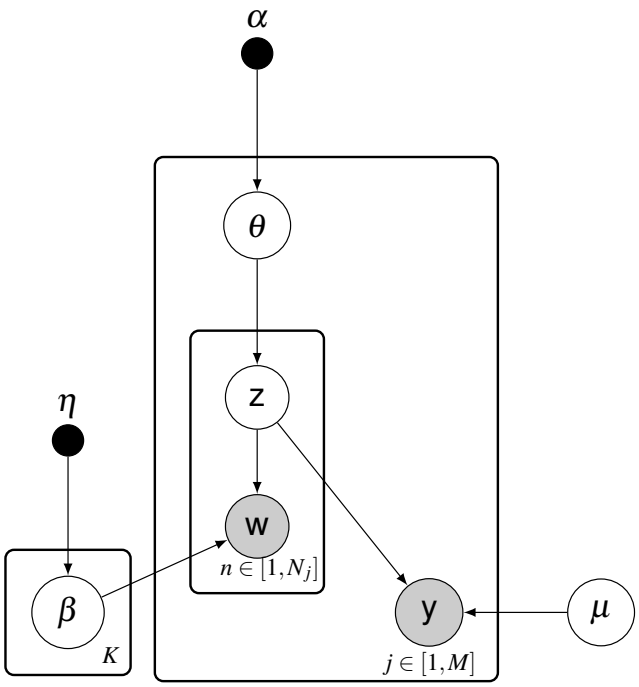
Topic Models for classification 1– LDAC



1

[1] Blei et.al. Latent Dirichlet allocation, JMLR 2003

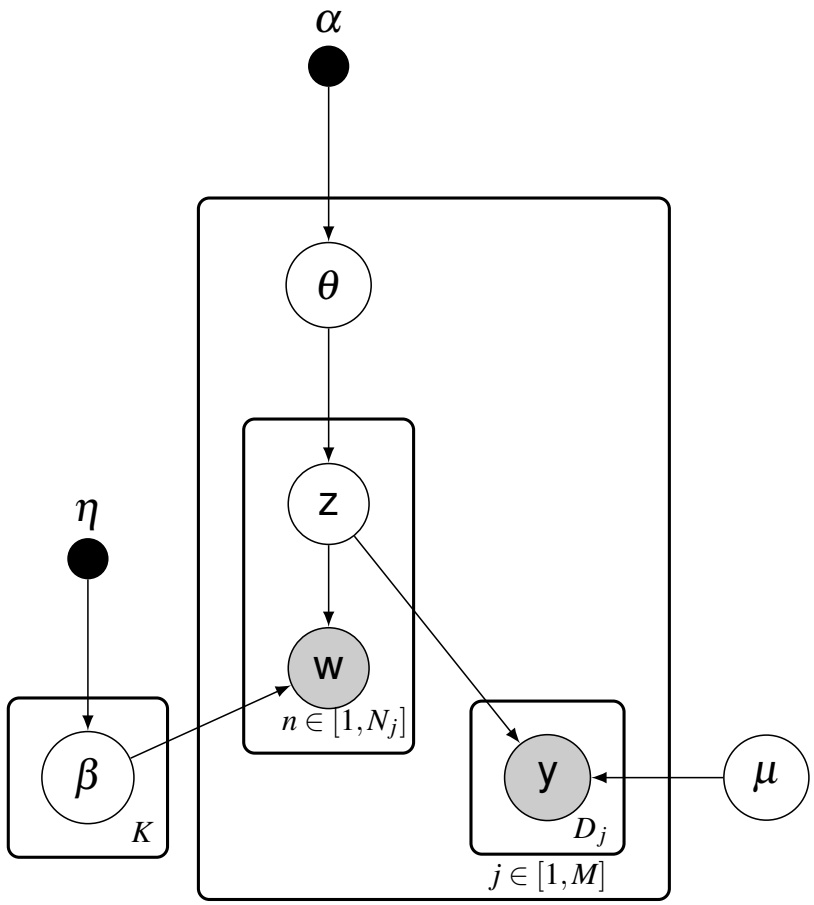
Topic Models for classification 2– SLDA



2

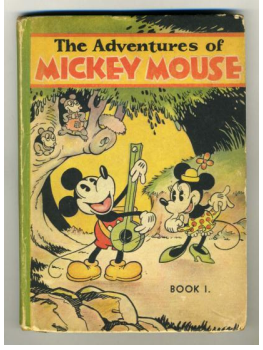
[2]Blei et.al. Supervised topic models, NIPS 2007

Topic Models for classification 3– PSLDA

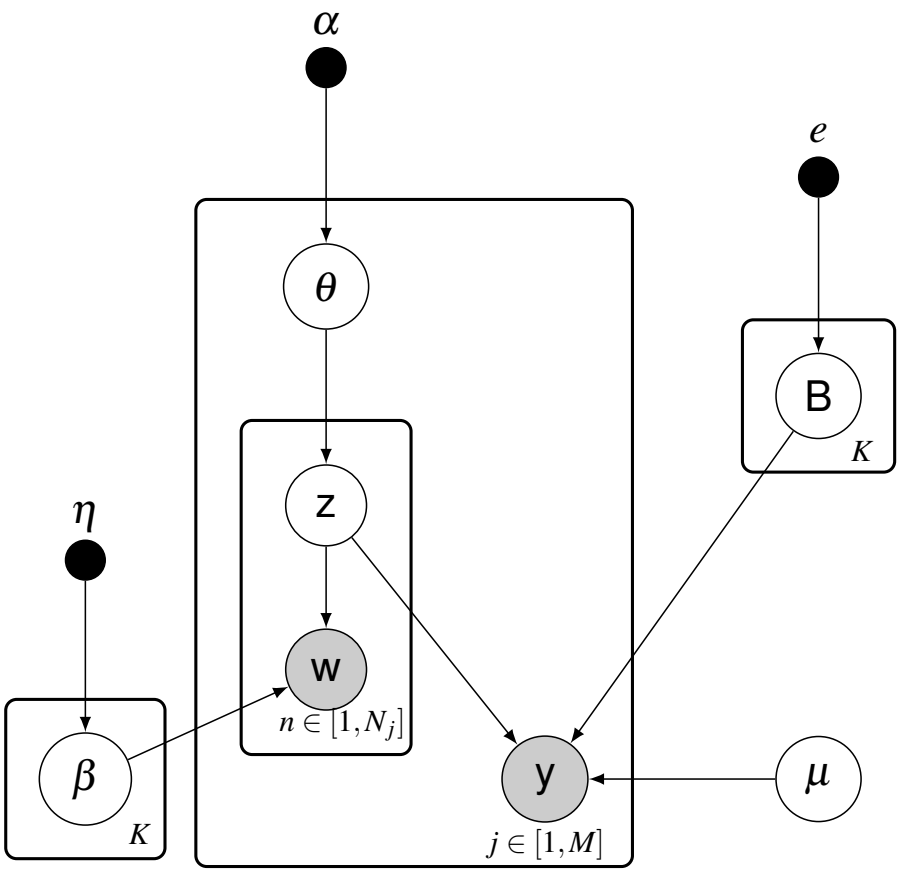


Topic Models for classification

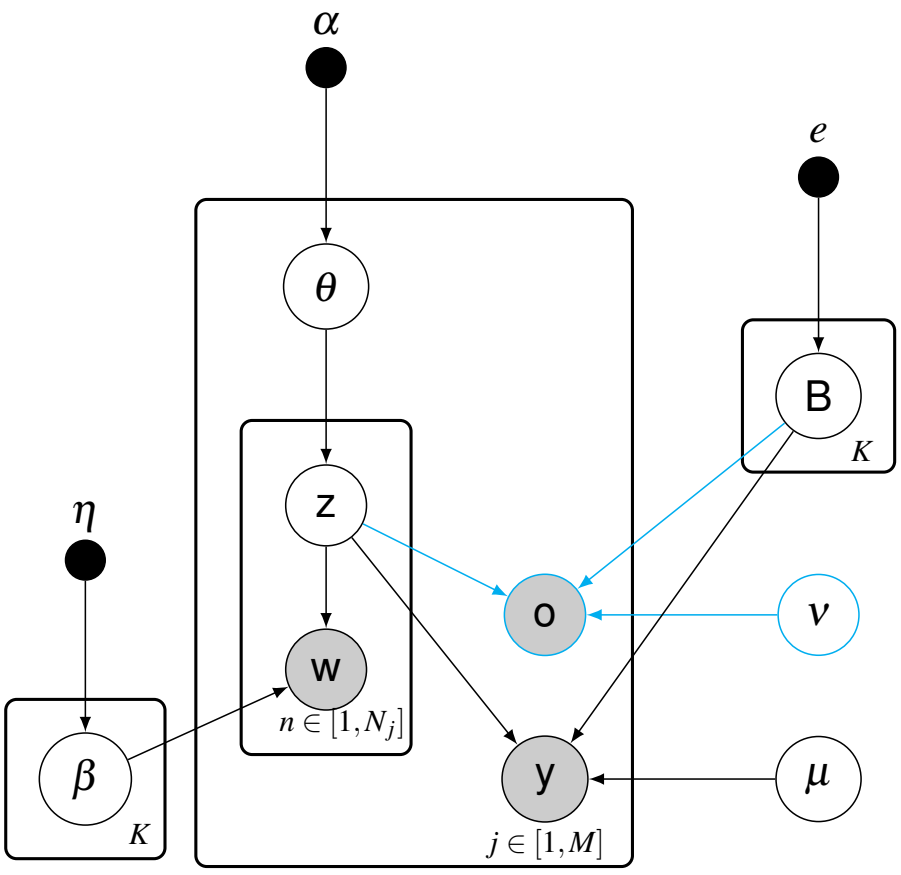




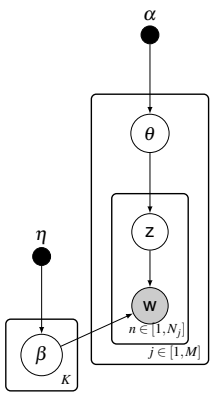
Topic Models for classification 4– NUFSLDA



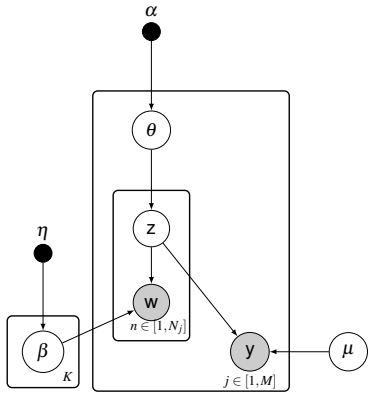
Topic Models for classification 5– NCFSLDA



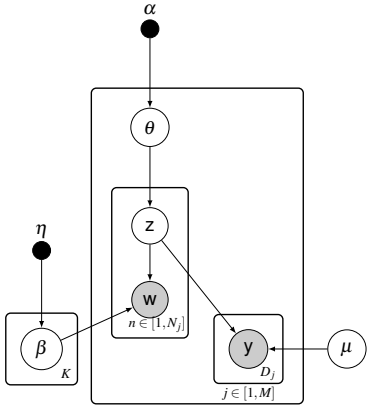
Topic Models for classification – Summation



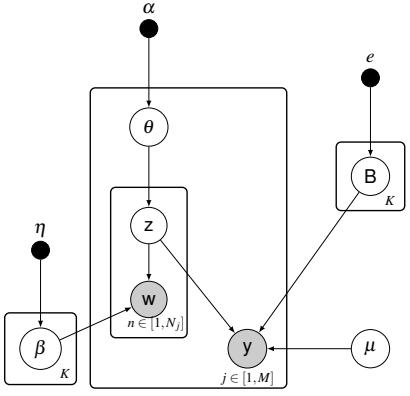
LDA



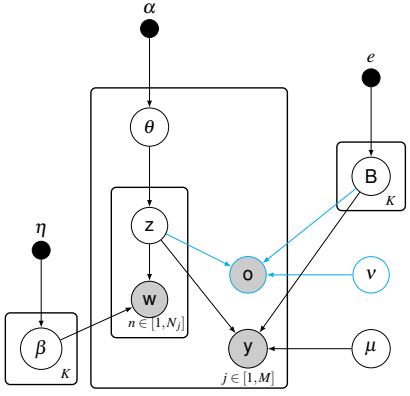
SLDA



P-SLDA

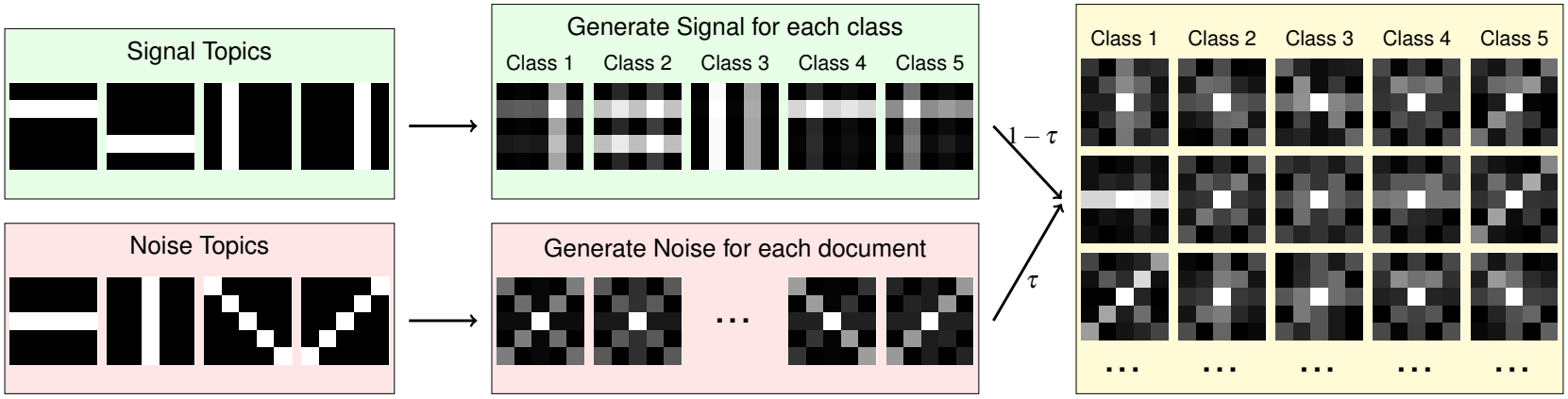


NUF-SLDA

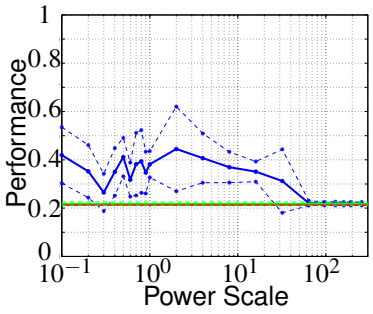


NCF-SLDA

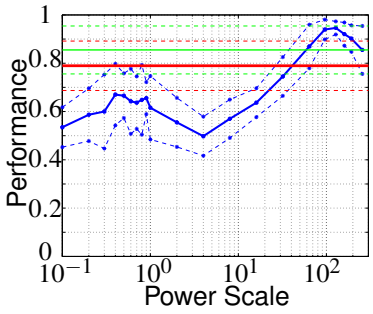
Experimental Evaluation –Synthetic Data



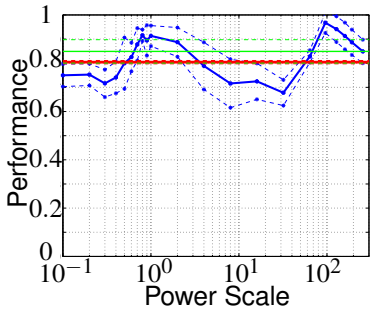
Experimental Evaluation –Synthetic Data



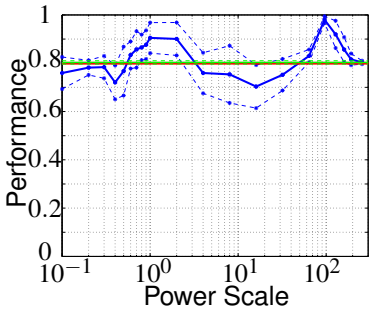
$K = 4,\ \tau = 0.9$



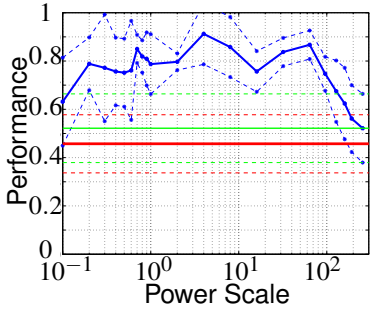
$K = 8, \tau = 0.9$



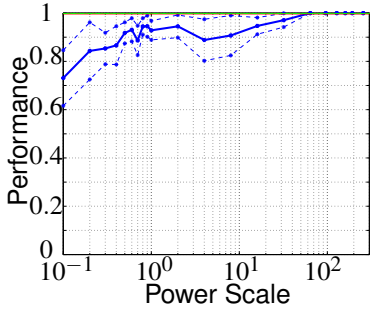
$K = 16, \tau = 0.9$



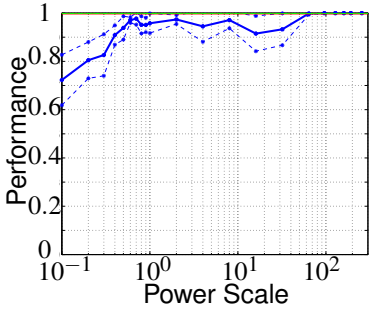
$K = 20, \tau = 0.9$



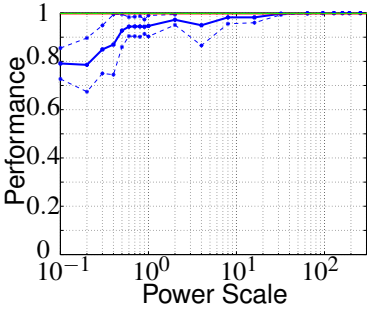
$K = 4,\ \tau = 0.7$



$K = 8, \tau = 0.7$



$K = 16, \tau = 0.7$

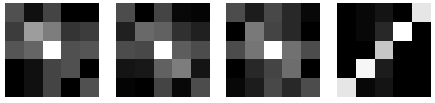


$K = 20, \tau = 0.7$

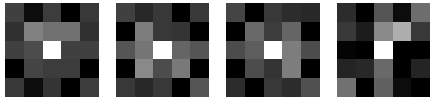
Experimental Evaluation –Synthetic Data

LDA

SLDA

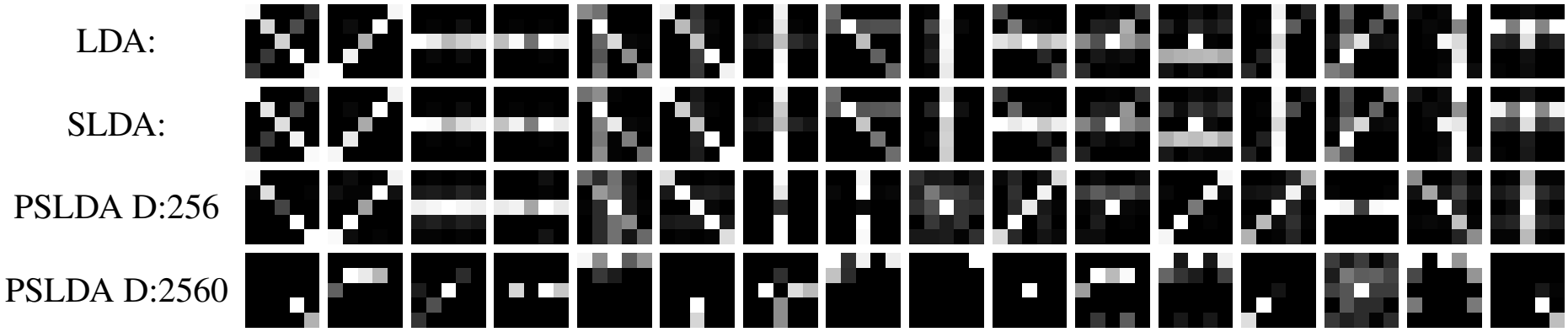


P-SLDA D=8



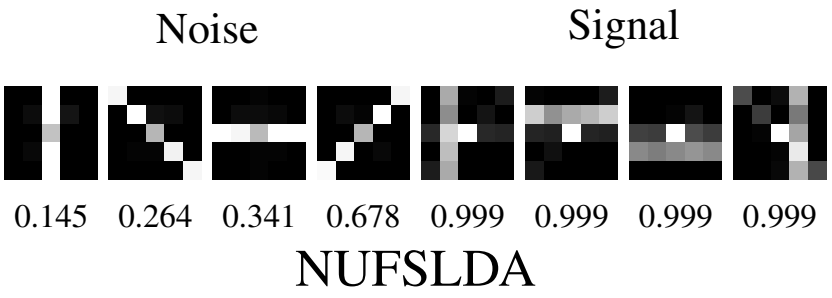
P-SLDA D: 256

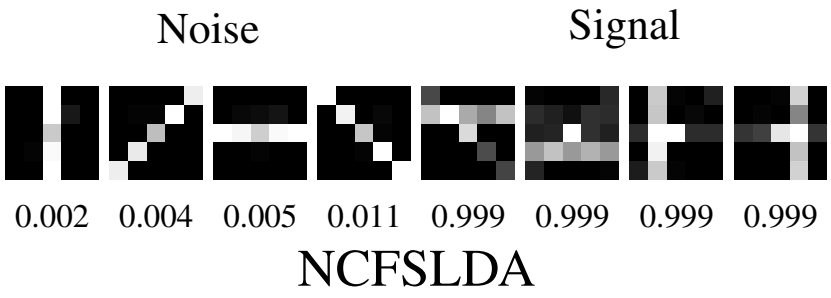
Experimental Evaluation –Synthetic Data



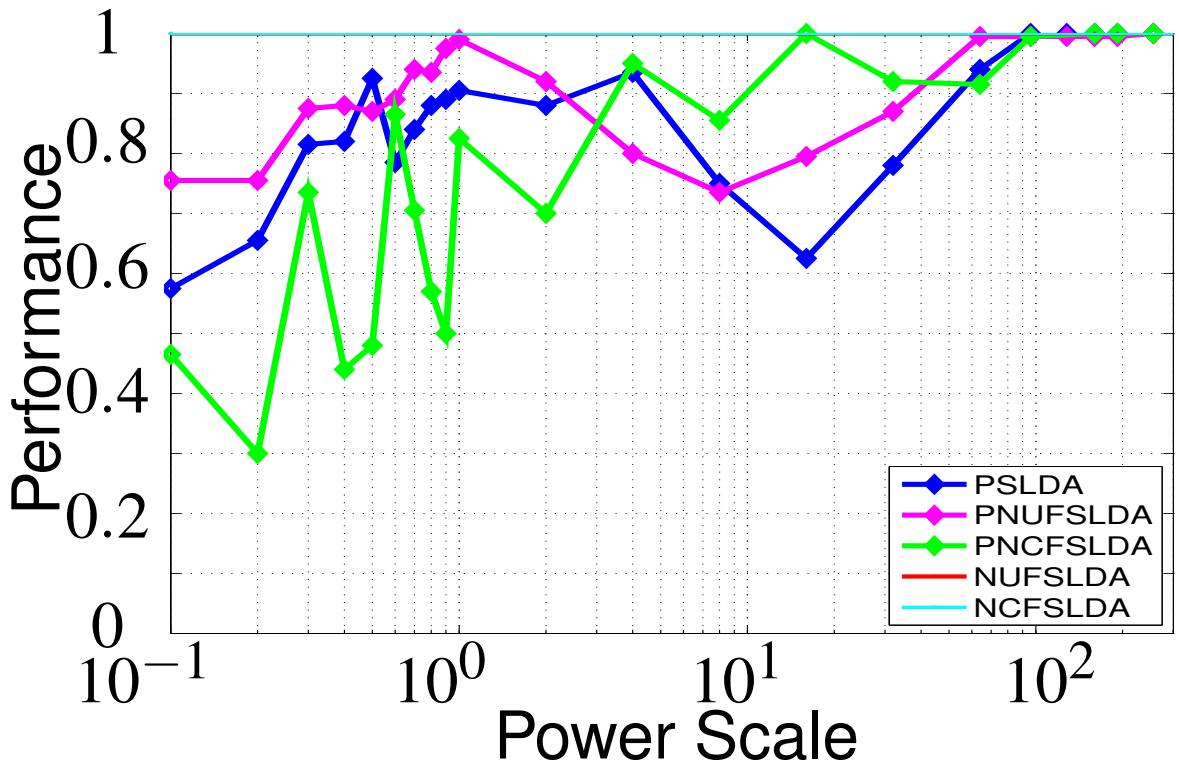
Experimental Evaluation –Synthetic Data

Learned topics by NUF-SLDA and NCF-SLDA

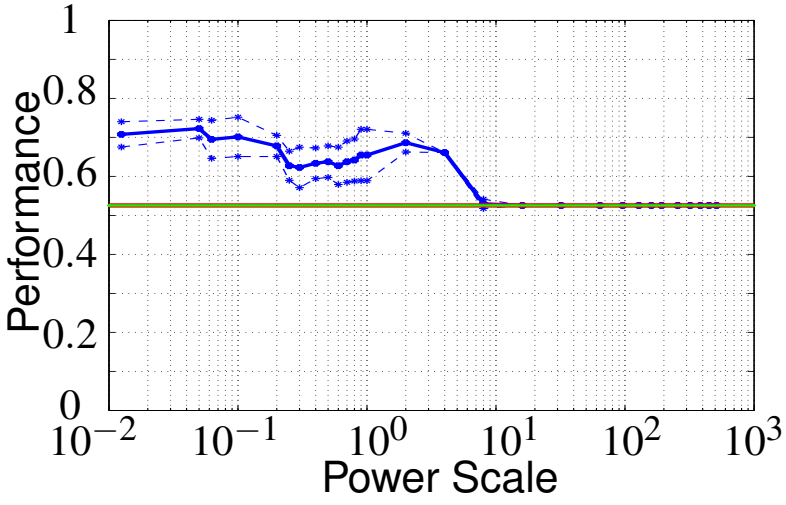




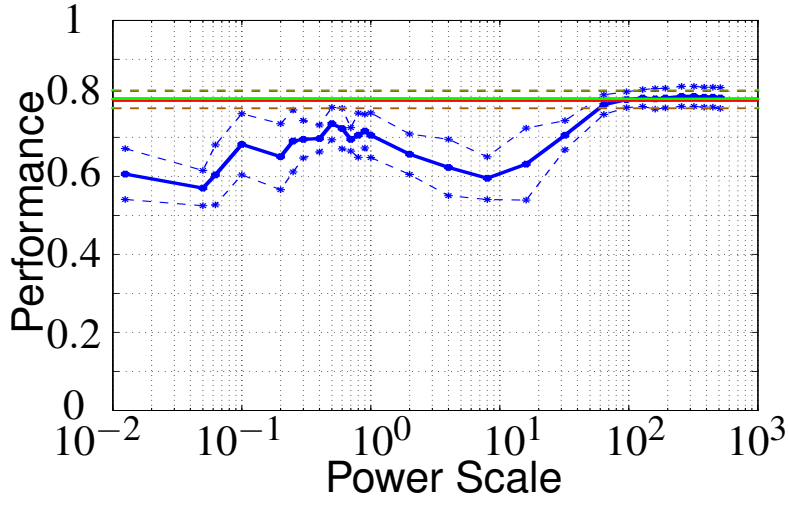
Experimental Evaluation –Synthetic Data



Experimental Evaluation –Action Data

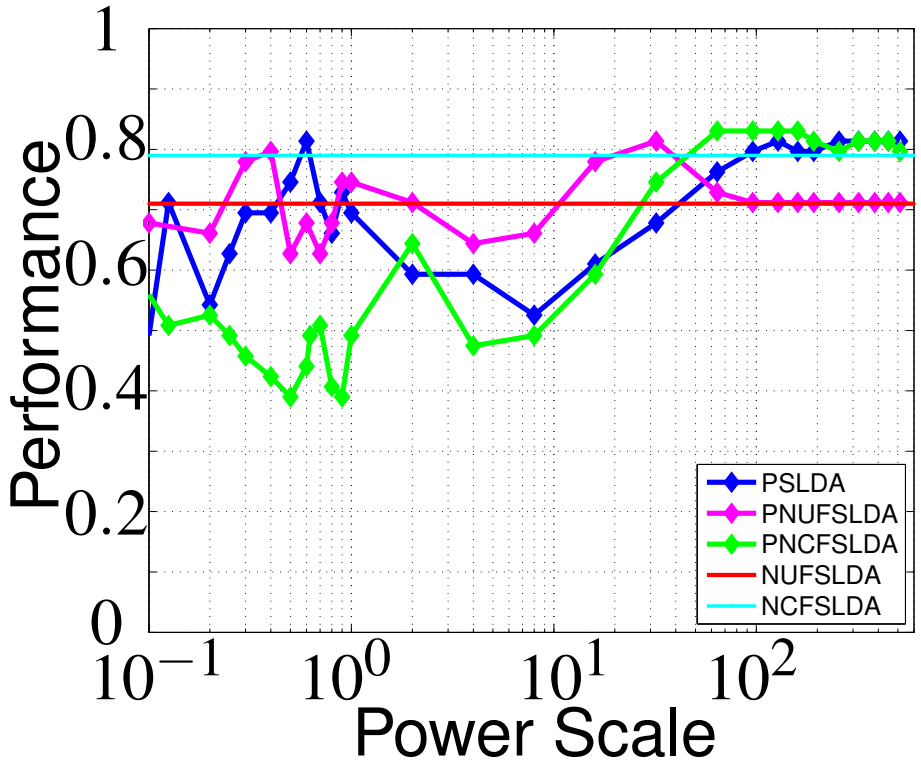


$K = 2$

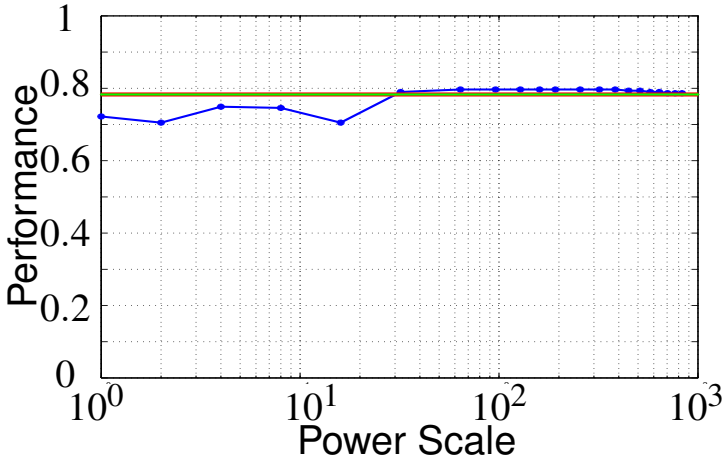


$K = 20$

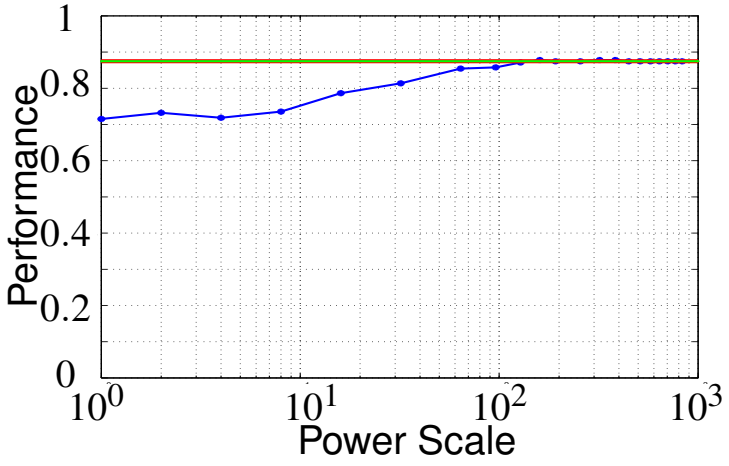
Experimental Evaluation –Action Data



Experimental Evaluation – Natural Scene Data

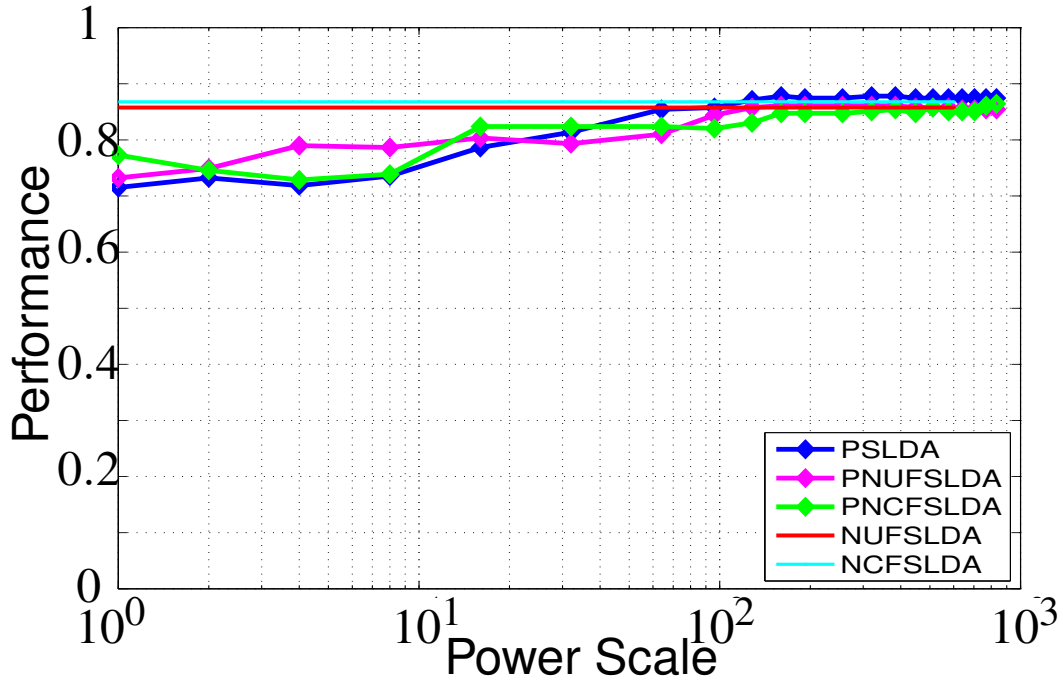


$K = 5$



$K = 30$

Experimental Evaluation – Natural Scene Data



Summary and Future work

1. Summary
    a) Using SLDA, the impact of supervision is limited on the learning of the latent space compared to the LDA
    b) Power SLDA (P-SLDA) can boost supervision
    c) Factorized Models can further enhance the performance
2. Future work
    a) Improve inference and explore more model framework for factorized models
    b) Analyze a wider range of supervised topic models

Thanks & Questions

Code Available:

http://www.csc.kth.se/~chengz/TopicModelCode.html

This is a C++ implementation of topic models with variational inference.

It has included LDA, supervised-LDA, HDP, supervised HDP, online HDP and online SHDP.

All the models that are used in this work will be available in this package.

Contact: chengz@kth.se