# Supplementary Material for "The Stochastic Topic Block Model for the Clustering of Vertices in Networks with Textual Edges" 

C. Bouveyron • P. Latouche • R. Zreik

Received: date / Accepted: date

1 Analysis of the Enron email network


Fig. 1 Model selection for STBM on the Enron data set. The values of ICL are divided by $10^{3}$.


Fig. 2 Reorganized adjacency matrix according to groups for STBM on the Enron data set.


Fig. 3 Specificity of a selection of words regarding the 5 found topics by STBM on the Enron data set.

## 2 Analysis of the Nips'15 co-authorship network



Fig. 4 Model selection for STBM on the NIPS co-authorship network. The values of ICL are divided by $10^{3}$.

Connexion probabilities between groups ( $\pi_{q}$ )

| 13 - | 0 | 0 | 0.01 | 0 | 0.28 | 0 | 0 | 0.14 | 0.02 | 0.05 | 0 | 0 | 0 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 12 - | 0 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0.59 | 0.07 | 0.2 | 0 | 1 | 0 |
| 11 - | 0 | 0 | 0.01 | 0 | 0 | 0 | 0 | 0.01 | 0 | 0.14 | 0.02 | 0 | 0 |
| $10-$ | 0 | 0 | 0.08 | 0 | 0.92 | 0 | 0 | 1 | 0.24 | 1 | 0.14 | 0.2 | 0.05 |
| 9 - | 0 | 0 | 0.03 | 0.01 | 0.53 | 0 | 0.01 | 0.27 | 0.1 | 0.24 | 0 | 0.07 | 0.02 |
| 8 - | 0 | 0.01 | 0.01 | 0.01 | 1 | 0.01 | 0 | 0.01 | 0.27 | 1 | 0.01 | 0.59 | 0.14 |
| 7 - | 0 | 0 | 0.01 | 0 | 0 | 0 | 0.01 | 0 | 0.01 | 0 | 0 | 0 | 0 |
| 6 - | 0 | 0 | 0.01 | 0 | 0 | 0.01 | 0 | 0.01 | 0 | 0 | 0 | 0 | 0 |
| 5 - | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0.53 | 0.92 | 0 | 0.2 | 0.28 |
| 4 - | 0 | 0 | 0.01 | 0.01 | 0 | 0 | 0 | 0.01 | 0.01 | 0 | 0 | 0 | 0 |
| $3-$ | 0 | 0.01 | 0 | 0.01 | 1 | 0.01 | 0.01 | 0.01 | 0.03 | 0.08 | 0.01 | 0 | 0.01 |
| 2 - | 0 | 0.01 | 0.01 | 0 | 0 | 0 | 0 | 0.01 | 0 | 0 | 0 | 0 | 0 |
| 1- | 0.01 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |

Fig. 5 Estimated matrix $\pi$ by STBM on the Nips co-authorship network.


Fig. 6 Reorganized adjacency matrix according to groups for STBM on the Nips co-authorship network.


Fig. 7 Specificity of a selection of words regarding the 5 found topics by STBM on the Nips co-authorship network.

