Supplementary materials

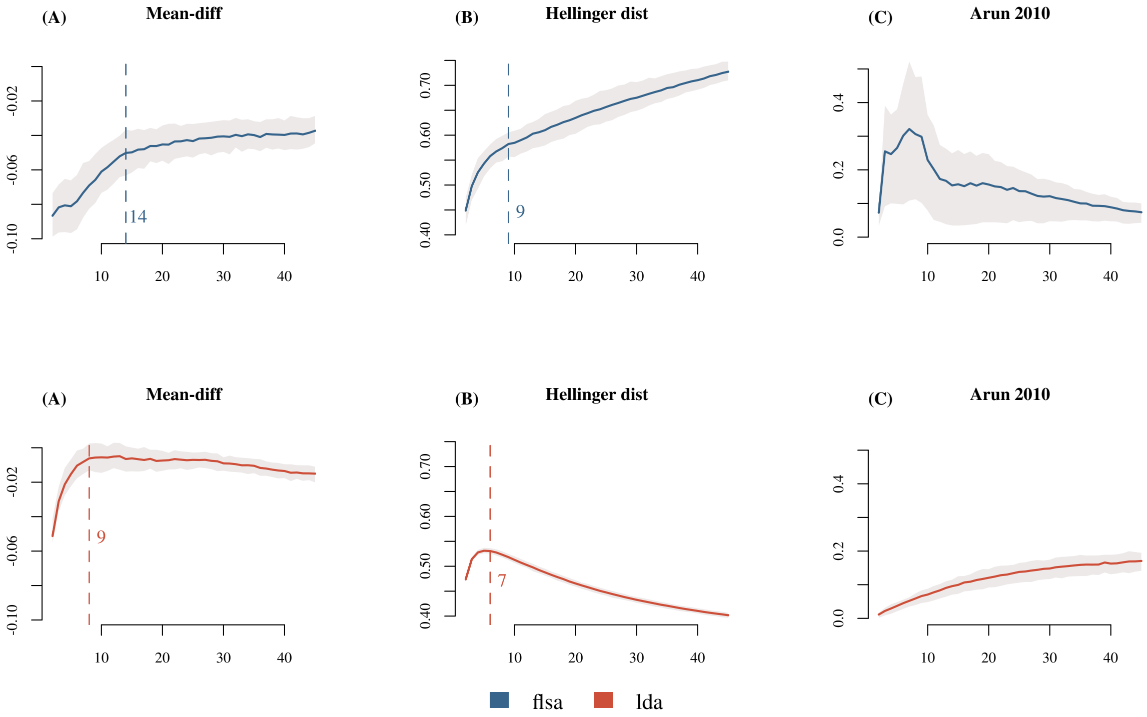

Fig. 1-SupMat Average quality measures for fLSA (blue curve) and LDA (red curve) as a function of the number of topics. Note that dashed vertical lines indicate the elbow of the curves, gray areas are the IQR ranges (i.e., $q_{0.75} - q_{0.25}$ tolerance intervals), and numbers indicate the corresponding elbow points (i.e., the optimal number of topics). All metrics have been computed using the top thirty topic words.

`Mean-diff`: coherence metric computed as in the `R` library `text2vec`.
`Hellinger distance`: Distance of each topic distribution from the so-called corpus distribution, which is computed as in the `R` library `topicmodels`.
`Arun 2010`: Topics similarity-based quality measure based on KL divergence

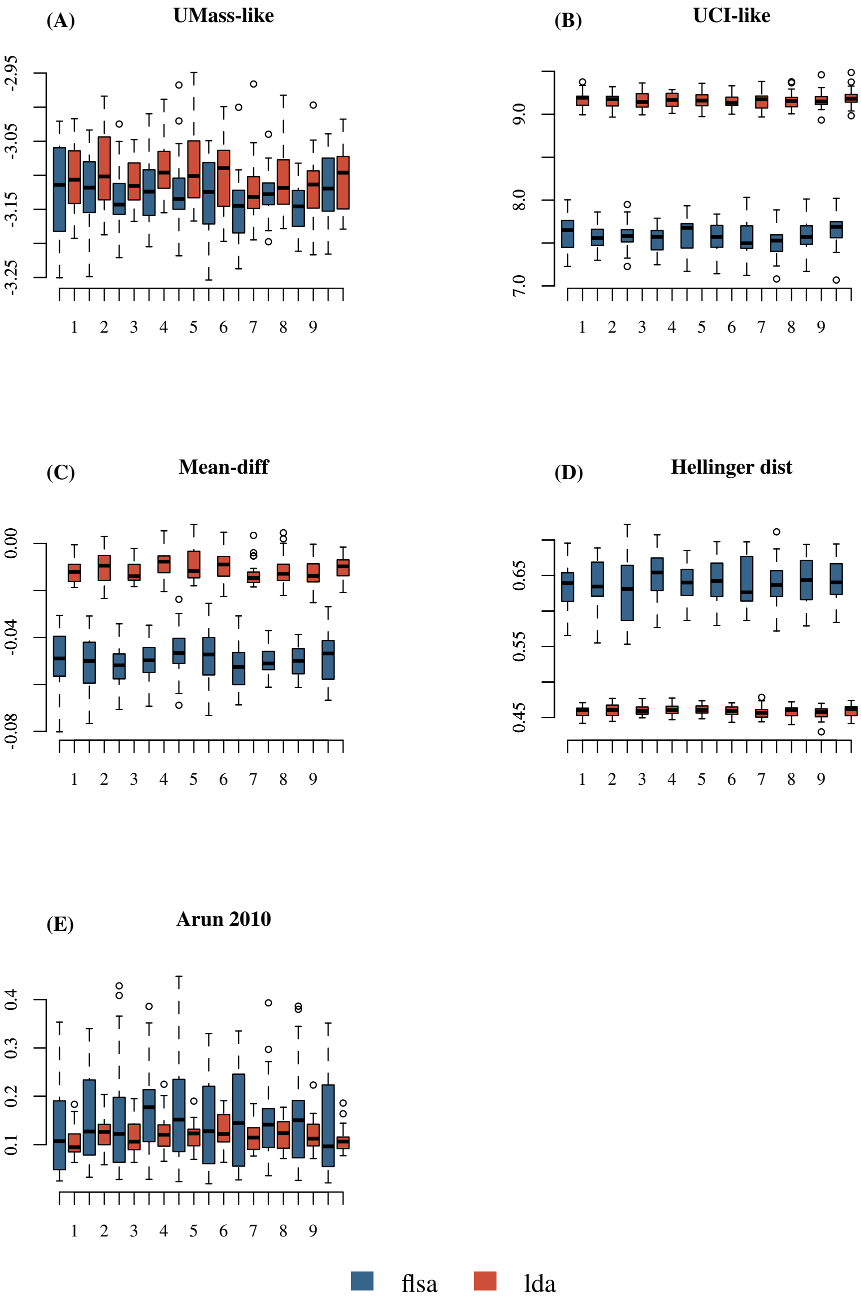

Fig. 2-SupMat Average Coherence and Quality measures for fLSA (blue boxes) and LDA (red boxes) as a function of the number of folds. All metrics have been computed using the top thirty topic words.

`UMass-like`: coherence metric computed as in the `R` library `text2vec`.
`UCI-like`: coherence metric computed as in the `R` library `text2vec`.
`Mean-diff`: coherence metric computed as in the `R` library `text2vec`.
`Hellinger distance`: Distance of each topic distribution from the so-called corpus distribution, which is computed as in the `R` library `topicmodels`.
`Arun 2010`: Topics similarity-based quality measure based on the KL divergence, which is computed as in the `R` library `ldatuning`.