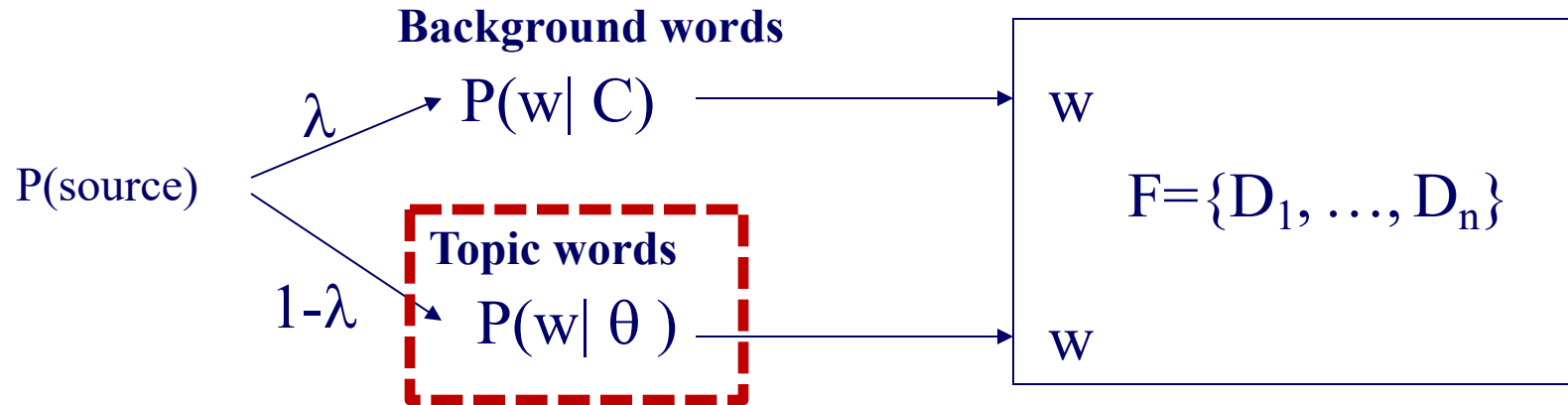


# $\theta_F$ Estimation Method I: Generative Mixture Model



$$\log p(F | \theta) = \sum_{D \in F} \prod_{w \in D} c(w; D) \log((1 - \lambda) p(w | \theta) + \lambda p(w | C))$$

**Maximum Likelihood**      $\theta_F = \underset{\theta}{\operatorname{argmax}} \log p(F | \theta)$

The learned topic model is called a “parsimonious language model” in [Hiemstra et al. 04]