ML Parameter Estimation

$$p_d(w) = \lambda_B p(w \mid \theta_B) + (1-\lambda_B)\sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)$$

$$\log p(d) = \sum_{w \in V} c(w,d) \log[\lambda_B p(w \mid \theta_B) + (1-\lambda_B)\sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)]$$

$$\log p(C \mid \Lambda) = \sum_{d \in C}\sum_{w \in V} c(w,d) \log[\lambda_B p(w \mid \theta_B) + (1-\lambda_B)\sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)]$$

Constrained Optimization: $\Lambda^* = \arg\max_{\Lambda} \mathrm{p}(\mathrm{C} \mid \Lambda)$

$$\forall \mathrm{j} \in [1,\mathrm{k}], \sum_{\mathrm{i}=1}^{\mathrm{M}} \mathrm{p}(\mathrm{w}_\mathrm{i} \mid \theta_\mathrm{j}) = 1 \qquad \forall \mathrm{d} \in \mathrm{C}, \sum_{\mathrm{j}=1}^{\mathrm{k}} \pi_{\mathrm{d,j}} = 1$$