$$egin{aligned} heta_w &= rac{\exp oldsymbol{\lambda}^ op \mathbf{f}(w)}{\sum_{w'} \exp oldsymbol{\lambda}^ op \mathbf{f}(w')} \ oldsymbol{\lambda}, \mathbf{x}) &= \sum_{oldsymbol{\lambda}} oldsymbol{\lambda}^ op \mathbf{f}(w) - \log \sum_{oldsymbol{\lambda}} oldsymbol{\lambda} \end{aligned}$$

$$\mathcal{L}(oldsymbol{\lambda}, \mathbf{x}) = \sum_{i=1}^{r-1} oldsymbol{\lambda}^{ op} \mathbf{f}(w) - \log \sum_{w'} \exp oldsymbol{\lambda}^{ op} \mathbf{f}(w')$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_k}(\boldsymbol{\lambda}, \mathbf{x}) = \sum_{i=1}^{|\mathbf{x}|} f_k(w) - \frac{\sum_{w'} f_k(w') \exp \boldsymbol{\lambda}^\top \mathbf{f}(w')}{\sum_{w''} \exp \boldsymbol{\lambda}^\top \mathbf{f}(w'')}$$

$$egin{aligned} rac{\partial \mathcal{L}}{\partial \lambda_k}(oldsymbol{\lambda}, \mathbf{x}) &= \sum_{i=1}^{n} f_k(w) - rac{\sum_{w'} J_k(w) \exp oldsymbol{\lambda}^{\top} \mathbf{f}(w')}{\sum_{w''} \exp oldsymbol{\lambda}^{\top} \mathbf{f}(w'')} \end{aligned}$$
 $= \sum_{i=1}^{|\mathbf{x}|} f_k(w) - \mathbb{E}_{p(w'; oldsymbol{\lambda})} f_k(w')$

$$\frac{\partial \mathcal{L}}{\partial \lambda_k}(\boldsymbol{\lambda}, \mathbf{x}) = \sum_{i=1}^{n} f_k(w) - \frac{\mathcal{L}_{w'} f_k(w)}{\sum_{w''} \exp \boldsymbol{\lambda}^{\top} \mathbf{f}(w')}$$

$$= \sum_{i=1}^{|\mathbf{x}|} f_k(w) - \mathbb{E}_{p(w'; \boldsymbol{\lambda})} f_k(w')$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_k}(\boldsymbol{\lambda}, \mathbf{x}) = 0$$