

$$\begin{aligned}
& \min_{\mathbf{w} \in \mathbb{R}^d, \mathbf{a} \in \mathbb{R}^K, \mathbf{b} \in \mathbb{R}^K} \max_{\mathbf{s} \in \mathbb{R}^K} \sum_{k=1}^K \mathbb{E}_{\mathbf{x}_i \sim \mathcal{S}_+^k} \left[(h_{\mathbf{w}}(\mathbf{x}_i; k) - a_k)^2 \right] + \mathbb{E}_{\mathbf{x}_j \sim \mathcal{S}_-^k} \left[(h_{\mathbf{w}}(\mathbf{x}_j; k) - b_k)^2 \right] \\
& \quad + s_k \left(\mathbb{E}_{\mathbf{x}_j \sim \mathcal{S}_-^k} [h_{\mathbf{w}}(\mathbf{x}_j; k)] - \mathbb{E}_{\mathbf{x}_i \sim \mathcal{S}_+^k} [h_{\mathbf{w}}(\mathbf{x}_i; k)] + c \right) - \ell^*(s_k),
\end{aligned}$$