

*Proof sketch of Corollary 4.6.*  $\mathbf{W} := [w_1^{(1)}, w_2^{(1)}, \dots, w_f^{(K)}]^\top \in \mathbb{R}^{fK \times d}$ , Note that by design,  $w_j^{(k)}$  and  $w_k$  always have equal norm. In this proof we consider the balanced case  $\mathbf{H} = [h_{1,1}, h_{1,2}, \dots, h_{1,n_1}, h_{2,1}, \dots, h_{K,n_K}] \in \mathbb{R}^{d \times N}$ , where  $N = \sum_{k=1}^K n_k$ . Denote  $\mathbf{Y} := [Y_1, \dots, Y_{fK}]$  the label of  $\mathbf{H}$  which consists of the columns  $Y_{f(k-1)+i_k} = [0, \dots, 0, 1, \dots, 1, 0, \dots, 0]^\top$ , where  $Y_{j,f(k-1)+i_k} = 1$  only when  $f(k-1) + 1 \leq j \leq f(k)$ .

$$\begin{aligned}
& \frac{1}{2fN} \|\mathbf{WH} - \mathbf{Y}\|_F^2 + \frac{\lambda \mathbf{w}_0}{2} \|\mathbf{W}_0\|_F^2 + \frac{\lambda_H}{2} \|\mathbf{H}\|_F^2 \\
&= \frac{1}{2f \sum_{k=1}^K n_k} \sum_{k=1}^K \sum_{i_k=1}^{n_k} \|\mathbf{W} \mathbf{h}_{k,i} - \mathbf{y}_k\|_2^2 + \frac{\lambda \mathbf{w}_0}{2} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \frac{\lambda_H}{2} \sum_{k=1}^K \sum_{i_k=1}^{n_k} \|\mathbf{h}_{k,i}\|_2^2 \\
&\stackrel{(a)}{\geq} \frac{1}{2f \sum_{k=1}^K n_k} \sum_{k=1}^K \sum_{j=1}^f n_k \frac{1}{n_k} \sum_{i_k=1}^{n_k} \left( \mathbf{w}_j^{(k)\top} \mathbf{h}_{k,i_k} - 1 \right)^2 + \frac{\lambda \mathbf{w}_0}{2} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \frac{\lambda_H}{2} \sum_{k=1}^K n_k \frac{1}{n_k} \sum_{i_k=1}^{n_k} \|\mathbf{h}_{k,i}\|_2^2 \\
&\stackrel{(b)}{\geq} \frac{1}{2f \sum_{k=1}^K n_k} \sum_{k=1}^K \sum_{j=1}^f n_k \left( \mathbf{w}_j^{(k)\top} \frac{1}{n_k} \sum_{i_k=1}^{n_k} \mathbf{h}_{k,i} - 1 \right)^2 + \frac{\lambda \mathbf{w}_0}{2} \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 + \frac{\lambda_H N}{2} \sum_{k=1}^K \frac{n_k}{N} \left\| \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbf{h}_{k,i_k} \right\|_2^2 \\
&\stackrel{(c)}{\geq} \frac{1}{2f \sum_{k=1}^K n_k} \sum_{k=1}^K n_k f \left( \frac{1}{f} \sum_{j=1}^f \mathbf{x}_j^{(k)} - 1 \right)^2 + \frac{\lambda \mathbf{w}_0}{2} K \left( \frac{1}{K} \sum_{k=1}^K \|\mathbf{w}_k\|_2 \right)^2 + \frac{\lambda_H N}{2} \left( \sum_{k=1}^K \frac{n_k}{N} \|\mathbf{h}_k\|_2 \right)^2 \\
&= \frac{1}{2 \sum_{k=1}^K n_k} \sum_{k=1}^K n_k \left( \frac{1}{f} \sum_{j=1}^f \mathbf{x}_j^{(k)} - 1 \right)^2 + \frac{\lambda \mathbf{w}_0}{2} K \left( \frac{1}{K} \sum_{k=1}^K \|\mathbf{w}_k\|_2 \right)^2 + \frac{\lambda_H N}{2} \left( \sum_{k=1}^K \frac{n_k}{N} \|\mathbf{h}_k\|_2 \right)^2 \\
&\stackrel{(d)}{\geq} \frac{1}{2} \left( \sum_{k=1}^K \frac{n_k}{\sum_{k=1}^K n_k} \mathbf{x}^{(k)} - 1 \right)^2 + \sqrt{KN \lambda_H \lambda \mathbf{w}_0} \|\mathbf{w}_k\|_2 \|\mathbf{h}_k\|_2
\end{aligned}$$

The inequality (a) follows from Jensen's Inequality and let  $\mathbf{w}^{(k')\top} \mathbf{h}_{k,i} = 0$  for all  $k' \neq k$  and  $i \in [n]$ . In (b) we used Jensen's inequality due to the strict convexity of  $(\cdot - 1)^2$  and  $\|\cdot\|^2$ .

Since all features in each class are identical, let us denote  $\mathbf{h}_k := \frac{1}{n} \sum_{i=1}^{n_k} \mathbf{h}_{k,i}$  and  $\mathbf{x}_j^{(k)} = \mathbf{w}_j^{(k)\top} \mathbf{h}_k$ . We get (c) by Jensen's inequality, which holds with equality iff

$$\begin{aligned}
& \mathbf{h}_{k,1} = \dots = \mathbf{h}_{k,n_k}, \quad \forall k \in [K] \\
& \mathbf{x}^{(k)} := \frac{1}{f} \sum_j \mathbf{x}_j^{(k)} = \mathbf{x}_1^{(k)} = \dots = \mathbf{x}_f^{(k)}, \quad \forall j \in [f] \\
& \|\mathbf{w}_1\|_2 = \dots = \|\mathbf{w}_K\|_2, \\
& \|\mathbf{h}_1\|_2 = \dots = \|\mathbf{h}_K\|_2, \\
& K \lambda \mathbf{w}_0 \|\mathbf{w}_k\|_2^2 = N \lambda_H \|\mathbf{h}_k\|_2^2
\end{aligned}$$

The last inequality degenerates to the equation (66) in the appendix of our paper. □