

## Errata

We discovered errors in (i) the value of  $\tau$  in equation (43), (ii) the value of  $\tau'$  in Algorithm 1, (iii) equation (59), and (iv) equations (84)-(85) in the Appendix.

(i) Before correcting (43), we provide the following Bellman equation for all  $\pi \in \Pi$  and for all  $(\delta, d)$ :

$$h_\pi(\delta, d) = \mathbb{E} \left[ \sum_{k=0}^{Z_\pi(\delta, d) + T_1(l_\pi(\delta, d)) - 1} \left( \text{err}_{\text{inference}}(\delta + k, d) - \bar{p}_\pi \right) \right] + \mathbb{E}[h_\pi(T_1(l_\pi(\delta, d)) + b_\pi(\delta, d), l_\pi(\delta, d))]. \quad (1)$$

Equation (1) is missing in the paper, but necessary. For a given state  $(\delta', d')$ , we consider

$$h_\pi(\delta', d') = 0, \forall \pi \in \Pi. \quad (2)$$

Now, we are ready to provide the correct equation for computing the average inference error  $\bar{p}_\pi$ . By using (1) and (2), we obtain:

$$\bar{p}_\pi = \frac{1}{\mathbb{E}[\tau]} \left( \mathbb{E} \left[ \sum_{k=0}^{\tau-1} \text{err}_{\text{inference}}(\delta' + k, d') \right] + \mathbb{E}[h_\pi(T_1(l_\pi(\delta', d')) + b_\pi(\delta', d'), l_\pi(\delta', d'))] \right), \quad (3)$$

where  $\tau = Z_\pi(\delta', d') + T_1(l_\pi(\delta', d'))$ .

(ii) The value of  $\tau'$  in Algorithm 1 is given by

$$\tau' \leftarrow Z_\pi(\delta, d) + T_1(l_\pi(\delta, d)). \quad (4)$$

(iii) The optimal feature length  $l_{j,\lambda}^*(t)$  is determined by

$$l_{j,\lambda}^*(t) = \arg \max_{l \in \mathbb{Z}: 0 \leq l \leq B_j} h_{j,\lambda}(\Delta(t) + 1, d(t)) - h_{j,\lambda}(\hat{b}_{j,\lambda}(l) + 1, l) - \lambda l. \quad (5)$$

(iv)  $V(\cdot)$  in (84)-(85) needs to be replaced by  $h(\cdot)$  as follows:

$$\begin{aligned}
h(\delta, d) &= \min_{\substack{Z \in \mathfrak{M} \\ l \in \mathbb{Z}: 1 \leq l \leq B \\ b \in \mathbb{Z}: 0 \leq b \leq B-l}} \left\{ \mathbb{E} \left[ \sum_{k=0}^{Z+T_{i+1}(l)-1} \text{err}_{\text{inference}}(\delta + k, d) \right] \right. \\
&\quad \left. - \mathbb{E}[Z + T_{i+1}(l)] \bar{p}_{\text{opt}} + \mathbb{E}[h(T_{i+1}(l) + b, l)] \right\} \\
&= \min_{\substack{Z \in \mathfrak{M} \\ l \in \mathbb{Z}: 1 \leq l \leq B \\ b \in \mathbb{Z}: 0 \leq b \leq B-l}} \left\{ \mathbb{E} \left[ \sum_{k=0}^{Z+T_1(l)-1} \text{err}_{\text{inference}}(\delta + k, d) \right] \right. \\
&\quad \left. - \mathbb{E}[Z + T_1(l)] \bar{p}_{\text{opt}} + \mathbb{E}[h(T_1(l) + b, l)] \right\}, \tag{6}
\end{aligned}$$

$$\begin{aligned}
h(\delta, d) &= \min_{\substack{l \in \mathbb{Z} \\ 1 \leq l \leq B}} \left\{ \mathbb{E} \left[ \sum_{k=0}^{Z_l(\delta, d)+T_1(l)-1} \left( \text{err}_{\text{inference}}(\delta + k, d) - \bar{p}_{\text{opt}} \right) \right] \right. \\
&\quad \left. + \min_{\substack{b \in \mathbb{Z} \\ 0 \leq b \leq B-l}} \mathbb{E}[h(T_1(l) + b, l)] \right\}. \tag{7}
\end{aligned}$$