Errata

We discovered errors in (i) the value of τ in equation (43), (ii) the value of τ' in Algorithm 1, (iii) equation (59),(iv) equations (84)-(85) in the Appendix, and (v) the scheduling policy space Π in Section II-C.

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

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

Equation (1) is missing in the paper, but necessary. For a given state (δ', 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:

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

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

(ii) The value of τ' 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

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

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

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

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

(v) Scheduling Policy Space Π:

Let $\pi = ((S_1, b_1, l_1), (S_2, b_2, l_2), \ldots)$ represent a scheduling policy. We focus on the class of *signal-agnostic* scheduling policies in which each decision is determined without using the knowledge of the signal value of the observed process. A scheduling policy π is said to be signal-agnostic, if the policy is independent of $\{(Y_t, X_t^l), t = 0, 1, 2, \ldots\}$. Let Π denote the set of all the causal scheduling policies that satisfy the following conditions: (i) the scheduling time S_i, the feature position b_i, and the feature length l_i are decided based on the current and the historical information available at the scheduler such that $1 \leq l_i \leq B$ and $0 \leq b_i \leq B - l_i$, (ii) the scheduler has access to the inference error function $\mathrm{err}_{\mathrm{inference}}(\cdot)$ and the distribution of $T_i(l)$ for each $l = 1, 2, \ldots, B$, and (iii) the scheduler does not have access to the realization of the process $\{(Y_t, X_t^l), t = 0, 1, 2, \ldots\}$.