Appendix
Proof of Theorem 1

We first recall the concentration inequality in [19], which extends Hoeffding's bound to possibly dependent bounded random variables.

Theorem 1 ([19, Theorem 1.8]). *Suppose that $X_1, \dots, X_n$ are random variables with $0 \le X_i \le 1$. Let $p = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[X_i]$, and fix ϑ with $np + 1 < \vartheta < n$. If $\epsilon_0 > 0$ satisfies*

$$\vartheta - 1 \; = \; np + np\epsilon_0,$$

then

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i \ge \vartheta\right) \; \le \; 2\exp\!\big(-n\, D(p(1+\epsilon_0)\,\|\,p)\big),$$

where $D(q\|p) \; = \; q\ln\frac{q}{p} + (1-q)\ln\frac{1-q}{1-p}$ denotes the Kullback–Leibler divergence.

Throughout this proof we use the following score-based distance between histories:

$$D_H(\mathcal{H}_i^t, \mathcal{H}_{i'}^t) \; := \; \max_{j\in\{i\}_{i=1}^{P}} \Big| S(\mathcal{H}_i^t, p_j) - S(\mathcal{H}_{i'}^t, p_j) \Big|. \tag{1}$$

To obtain the most conservative estimate, we evaluate the score $S(\mathcal{H}_i^t, p_j)$ using uniform temporal weights. This corresponds to the worst case, since no slot is favored and temporal differences cannot increase the separation between histories. Therefore, for the prototype score, we use uniform temporal weights, i.e., $w_{t-1-l} \; = \; \frac{1}{W}, \quad l = t-W, \dots, t-1$, so that

$$S(\mathcal{H}, p_j) = \frac{1}{W}\sum_{l=t-W}^{t-1} d\big(l, p_j; \mathcal{H}\big), \qquad d(\cdot) \in [0,1].$$

Let $\mathcal{H}_i^t$ and $\mathcal{H}_{i'}^t$ be two (random) histories such that $D_H(\mathcal{H}_i^t, \mathcal{H}_{i'}^t) \ge \delta$. For each history, the selected prototype index is $\kappa_i^t \in \arg\max_{j\in\{i\}_{i=1}^{P}} S(\mathcal{H}_i^t, p_j)$, and we set

$$k \triangleq \kappa_i^t, \qquad k' \triangleq \kappa_{i'}^t.$$

By (1), there exist indices

$$j^+ \in \arg\max_j \Big(S(\mathcal{H}_i^t, p_j) - S(\mathcal{H}_{i'}^t, p_j)\Big),$$

$$j^- \in \arg\max_j \Big(S(\mathcal{H}_{i'}^t, p_j) - S(\mathcal{H}_i^t, p_j)\Big),$$

such that

$$\begin{aligned} S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+}) \; &\ge \; \delta, \\ S(\mathcal{H}_{i'}^t, p_{j^-}) - S(\mathcal{H}_i^t, p_{j^-}) \; &\ge \; \delta. \end{aligned} \tag{2}$$

Lemma 1 (Alignment in expectation for the witness prototypes). *Let j^+, j^- be as above. Then*

$$\frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}\Big[d\big(l, p_{j^+}; \mathcal{H}_i^t\big) - d\big(l, p_{j^+}; \mathcal{H}_{i'}^t\big)\Big] \; \ge \; \delta, \tag{3}$$

and

$$\frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}\Big[d\big(l, p_{j^-}; \mathcal{H}_{i'}^t\big) - d\big(l, p_{j^-}; \mathcal{H}_i^t\big)\Big] \; \ge \; \delta. \tag{4}$$

Proof. By the definition of $S(\cdot, p_j)$ and linearity of expectation,

$$\mathbb{E}\Big[S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+})\Big] = \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}\Big[d\big(l, p_{j^+}; \mathcal{H}_i^t\big) - d\big(l, p_{j^+}; \mathcal{H}_{i'}^t\big)\Big].$$

Combining with the first inequality in (2) gives (3). The argument for j^- is identical, using the second inequality in (2). □

We now show that this witness alignment suffices to obtain the desired probability bound. We define, for $l = t-W, \dots, t-1$,

$$Y_l \; = \; d\big(l, p_{j^+}; \mathcal{H}_i^t\big) - d\big(l, p_{j^+}; \mathcal{H}_{i'}^t\big).$$

By construction of $d(\cdot)$ and the weights w_a, w_M, w_m, w_ρ, each $d(\cdot)$ lies in $[0,1]$, hence $-1 \le Y_l \le 1, \quad l = t-W, \dots, t-1$. Moreover,

$$S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+}) = \frac{1}{W}\sum_{l=t-W}^{t-1} Y_l.$$

Introduce the rescaled variables

$$X_l = \frac{1-Y_l}{2}, \qquad l = t-W, \dots, t-1,$$

so that $0 \le X_l \le 1$ and $Y_l = 1 - 2X_l$. We also have

$$S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+}) = \frac{1}{W}\sum_{l=t-W}^{t-1} Y_l = 1 - \frac{2}{W}\sum_{l=t-W}^{t-1} X_l.$$

Let

$$p \; = \; \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}[X_l].$$

Taking expectations and using (3), we obtain

$$\begin{aligned} \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}[Y_l] &= \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}\Big[d(l, p_{j^+}; \mathcal{H}_i^t) - d(l, p_{j^+}; \mathcal{H}_{i'}^t)\Big] \\ &\ge \delta, \end{aligned}$$

so that

$$\begin{aligned} p = \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}[X_l] &= \frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}\Big[\frac{1-Y_l}{2}\Big] \\ &= \frac{1}{2} - \frac{1}{2W}\sum_{l=t-W}^{t-1} \mathbb{E}[Y_l] \\ &\le \frac{1}{2} - \frac{\delta}{2}. \end{aligned}$$

We now set up Theorem 1 for the sequence $X_{t-W}^{(+)}, \dots, X_{t-1}^{(+)}$. For any $\delta \in (0,1)$ and $W > \frac{2}{1-\delta}$, we can choose a parameter λ satisfying

$$\frac{1}{W\delta} < \lambda < \frac{1}{2\delta} - \frac{1}{2}.$$

Define the threshold

$$\vartheta_+ \;=\; Wp + \lambda W\delta, \qquad q \;=\; \frac{\vartheta_+}{W} = p + \lambda\delta.$$

We first verify that the conditions of Theorem 1 hold. Using $\lambda > \frac{1}{W\delta}$,

$$\vartheta_+ - Wp = \lambda W\delta > \frac{1}{W\delta}\,W\delta = 1,$$

so $Wp + 1 < \vartheta_+$. On the other hand,

$$q = p + \lambda\delta \le \left(\frac{1}{2} - \frac{\delta}{2}\right) + \lambda\delta.$$

Using $\lambda < \frac{1}{2\delta}$, we have

$$q < \frac{1}{2} - \frac{\delta}{2} + \frac{\delta}{2} = \frac{1}{2},$$

hence $\vartheta_+ = Wq < W$. Therefore $Wp + 1 < \vartheta_+ < W$, and the hypothesis of Theorem 1 is satisfied for the sequence $X^{(+)}_{t-W}, \ldots, X^{(+)}_{t-1}$ and threshold ϑ_+.

Applying Theorem 1,

$$\mathbb{P}\left(\sum_{l=t-W}^{t-1} X_l \ge \vartheta_+\right) \;\le\; 2\exp\bigl(-W\,D(q\|p)\bigr).$$

Using the standard quadratic lower bound

$$D(q\|p) \;\ge\; \frac{(q-p)^2}{2p(1-p)},$$

and since $p(1-p) \le 1/4$, we obtain

$$D(q\|p) \;\ge\; 2(q-p)^2 = 2\lambda^2\delta^2.$$

Hence

$$\mathbb{P}\left(\sum_{l=t-W}^{t-1} X_l \ge \vartheta_+\right) \;\le\; 2\exp\bigl(-2W\lambda^2\delta^2\bigr).$$

Observe that

$$\sum_{l=t-W}^{t-1} Y_l = W - 2\sum_{l=t-W}^{t-1} X_l.$$

The event $\{\sum_l Y_l \le 0\}$ is equivalent to $\{\sum_l X_l \ge W/2\}$. Moreover, since $q < 1/2$, we have $\vartheta_+ = Wq < W/2$, so $\{\sum_l X_l \ge W/2\} \subseteq \{\sum_l X_l \ge \vartheta_+\}$. Therefore

$$\begin{aligned}
\mathbb{P}\left(\sum_{l=t-W}^{t-1} Y_l \le 0\right) &= \mathbb{P}\left(\sum_{l=t-W}^{t-1} X_l \ge \frac{W}{2}\right)\\
&\le \mathbb{P}\left(\sum_{l=t-W}^{t-1} X_l \ge \vartheta_+\right)\\
&\le 2\exp\bigl(-2W\lambda^2\delta^2\bigr).
\end{aligned}$$

Equivalently,

$$\mathbb{P}\left(\sum_{l=t-W}^{t-1} Y_l > 0\right) \;\ge\; 1 - 2\exp\bigl(-2W\lambda^2\delta^2\bigr).$$

Recalling that

$$S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+}) = \frac{1}{W}\sum_{l=t-W}^{t-1} Y_l,$$

we conclude that

$$\mathbb{P}\bigl(S(\mathcal{H}_i^t, p_{j^+}) > S(\mathcal{H}_{i'}^t, p_{j^+})\bigr) \;\ge\; 1 - 2\exp\bigl(-2W\lambda^2\delta^2\bigr). \tag{5}$$

An identical argument applies to the witness prototype p_{j^-} associated with history $\mathcal{H}_{i'}^t$. Define

$$Y_l^{(-)} = d\bigl(l, p_{j^-}; \mathcal{H}_{i'}^t\bigr) - d\bigl(l, p_{j^-}; \mathcal{H}_i^t\bigr),$$

and $X_l^{(-)} = (1 - Y_l^{(-)})/2$. Using the alignment condition (4) from Lemma 1, we obtain

$$\frac{1}{W}\sum_{l=t-W}^{t-1} \mathbb{E}[Y_l^{(-)}] \;\ge\; \delta,$$

and the same application of Theorem 1 yields, for the same λ,

$$\mathbb{P}\bigl(S(\mathcal{H}_{i'}^t, p_{j^-}) > S(\mathcal{H}_i^t, p_{j^-})\bigr) \;\ge\; 1 - 2\exp\bigl(-2W\lambda^2\delta^2\bigr). \tag{6}$$

Define the events

$$\begin{aligned}
\mathcal{E}_1 &= \bigl\{S(\mathcal{H}_i^t, p_{j^+}) > S(\mathcal{H}_{i'}^t, p_{j^+})\bigr\},\\
\mathcal{E}_2 &= \bigl\{S(\mathcal{H}_{i'}^t, p_{j^-}) > S(\mathcal{H}_i^t, p_{j^-})\bigr\}.
\end{aligned}$$

From (5) and (6) and the union bound,

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) &= 1 - \mathbb{P}(\mathcal{E}_1^c \cup \mathcal{E}_2^c)\\
&\ge 1 - \mathbb{P}(\mathcal{E}_1^c) - \mathbb{P}(\mathcal{E}_2^c)\\
&\ge 1 - 4\exp\bigl(-2W\lambda^2\delta^2\bigr).
\end{aligned}$$

On the event $\mathcal{E}_1 \cap \mathcal{E}_2$, the two histories exhibit strictly opposite directed score gaps at the witness prototypes p_{j^+} and p_{j^-}:

$$S(\mathcal{H}_i^t, p_{j^+}) > S(\mathcal{H}_{i'}^t, p_{j^+}), \qquad S(\mathcal{H}_{i'}^t, p_{j^-}) > S(\mathcal{H}_i^t, p_{j^-}).$$

We now show that this forces the argmax indices to differ.

Suppose, for contradiction, that $\kappa_i^t = \kappa_{i'}^t$, and denote the common index by k. By maximality of p_k for each history,

$$S(\mathcal{H}_i^t, p_k) \ge S(\mathcal{H}_i^t, p_{j^-}), \qquad S(\mathcal{H}_{i'}^t, p_k) \ge S(\mathcal{H}_{i'}^t, p_{j^+}).$$

Combining these with the strict inequalities in $\mathcal{E}_1 \cap \mathcal{E}_2$ and the witness gaps (2) yields an inconsistency (loosely speaking, p_k cannot simultaneously dominate both witnesses while the witnesses prefer opposite histories by at least δ). Formally, from (2),

$$S(\mathcal{H}_i^t, p_{j^+}) \ge S(\mathcal{H}_{i'}^t, p_{j^+}) + \delta, \qquad S(\mathcal{H}_{i'}^t, p_{j^-}) \ge S(\mathcal{H}_i^t, p_{j^-}) + \delta,$$

and combining with the maximality inequalities leads to

$$S(\mathcal{H}_i^t, p_k) - S(\mathcal{H}_{i'}^t, p_k) \;\ge\; S(\mathcal{H}_i^t, p_{j^+}) - S(\mathcal{H}_{i'}^t, p_{j^+}) \;\ge\; \delta$$

and

$$S(\mathcal{H}_{i'}^t, p_k) - S(\mathcal{H}_i^t, p_k) \;\ge\; S(\mathcal{H}_{i'}^t, p_{j^-}) - S(\mathcal{H}_i^t, p_{j^-}) \;\ge\; \delta,$$

which is impossible. Hence on $\mathcal{E}_1 \cap \mathcal{E}_2$ we must have $\kappa_i^t \neq \kappa_{i'}^t$:

$$\mathcal{E}_1 \cap \mathcal{E}_2 \;\subseteq\; \bigl\{\kappa_i^t \neq \kappa_{i'}^t\bigr\}.$$

Consequently,

$$\mathbb{P}\bigl(\kappa_i^t \neq \kappa_{i'}^t\bigr) \;\geq\; \mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \;\geq\; 1 - 4\exp\bigl(-2W\lambda^2\delta^2\bigr)\,.$$

Finally, we can absorb the prefactor 4 into the exponential term. That is, there exists $c' > 0$ (depending only on λ) such that

$$\mathbb{P}\bigl(\kappa_i^t \neq \kappa_{i'}^t\bigr) \;\geq\; 1 - \exp\bigl(-c'W\delta^2\bigr).$$

This establishes the claimed bound and completes the proof of Theorem 1.

References (Appendix)

[19] C. Pelekis and J. Ramon, "Hoeffding's inequality for sums of weakly dependent random variables," 2015. [Online]. Available: https://arxiv.org/abs/1507.06871