A Proofs

Proof of Theorem 1. We will prove $\Theta_t^\star = \Theta_t$ by induction for all $0 \leq t$.

First, we assume, w.l.o.g., that the data sets $X_1, \cdots, X_t$ belong to a class c ($y_1 = \cdots = y_t = c$). So that, $\Theta_t^\star$ is a tuple $\langle(\vec{\mu}^\star_{(c,t)}, \mathbf{\Sigma}^\star_{(c,t)}) \mid c \in \{1, \ldots, C\}\rangle$, where:

$$\begin{aligned}
\vec{\mu}^\star_{(c,t)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} \vec{x} + \sum\limits_{\vec{x}\in X_1} \vec{x} + \cdots + \sum\limits_{\vec{x}\in X_t} \vec{x}}{|\mathcal{I}_c| + |X_1| + \cdots + |X_t|};\\
\mathbf{\Sigma}^\star_{(c,t)} &= \frac{1}{|\mathcal{I}_c| + |X_1| + \cdots + |X_t| - 1}\\
&\quad \times \Big[\sum_{\vec{x}\in\mathcal{I}_c} (\vec{x} - \vec{\mu}^\star_{(c,t)})(\vec{x} - \vec{\mu}^\star_{(c,t)})^T\\
&\quad + \sum_{\vec{x}\in X_1} (\vec{x} - \vec{\mu}^\star_{(c,t)})(\vec{x} - \vec{\mu}^\star_{(c,t)})^T\\
&\quad + \cdots + \sum_{\vec{x}\in X_t} (\vec{x} - \vec{\mu}^\star_{(c,t)})(\vec{x} - \vec{\mu}^\star_{(c,t)})^T \Big]'
\end{aligned}$$

and Θ_t is a tuple $\langle(\vec{\mu}_{(c,t)}, \mathbf{\Sigma}_{(c,t)}) \mid c \in \{1, \ldots, C\}\rangle$, where:

$$\begin{aligned}
\vec{\mu}_{(c,t)} &= \begin{cases} \vec{\mu}_c & \text{if } t = 0\\ \frac{N_{(c,t-1)}\vec{\mu}_{(c,t-1)} + |X_t|\vec{\mu}_{[X_t]}}{N_{(c,t-1)} + |X_t|} & \text{otherwise;} \end{cases}\\
\mathbf{\Sigma}_{(c,t)} &= \begin{cases} \mathbf{\Sigma}_c & \text{if } t = 0\\ \frac{(N_{(c,t-1)}-1)\mathbf{\Sigma}_{(c,t-1)} + (|X_t|-1)\mathbf{\Sigma}_{X_t} + \boldsymbol{\sigma}_{(c,t)}}{N_{(c,t-1)} + |X_t| - 1} & \text{otherwise.} \end{cases}
\end{aligned}$$

Base Case: When $t = 0$,

$$\begin{aligned}
\Theta_0^\star &= \langle(\vec{\mu}^\star_{(c,0)}, \mathbf{\Sigma}^\star_{(c,0)}) \mid c \in \{1, \ldots, C\}\rangle;\\
\vec{\mu}^\star_{(c,0)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} \vec{x}}{|\mathcal{I}_c|} = \vec{\mu}_c = \vec{\mu}_{(c,0)};\\
\mathbf{\Sigma}^\star_{(c,0)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} (\vec{x} - \vec{\mu}^\star_{(c,0)})(\vec{x} - \vec{\mu}^\star_{(c,0)})^T}{|\mathcal{I}_c| - 1}\\
&= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} (\vec{x} - \vec{\mu}_c)(\vec{x} - \vec{\mu}_c)^T}{|\mathcal{I}_c| - 1} = \mathbf{\Sigma}_c = \mathbf{\Sigma}_{(c,0)}.
\end{aligned}$$

Since $\vec{\mu}^\star_{(c,0)} = \vec{\mu}_{(c,0)}$ and $\mathbf{\Sigma}^\star_{(c,0)} = \mathbf{\Sigma}_{(c,0)}$, then

$$\Theta_0^\star = \langle(\vec{\mu}_{(c,0)}, \mathbf{\Sigma}_{(c,0)}) \mid c \in \{1, \ldots, C\}\rangle = \Theta_0,$$

so the base case holds.

Inductive step: Suppose that $\Theta_k^\star = \Theta_k$ holds for $0 \leq k$. We will show that $\Theta_{k+1}^\star = \Theta_{k+1}$.

$$\Theta_{k+1}^\star = \langle(\vec{\mu}^\star_{(c,k+1)}, \mathbf{\Sigma}^\star_{(c,k+1)}) \mid c \in \{1, \ldots, C\}\rangle,$$

starting with the mean vector:

$$\begin{aligned}
\vec{\mu}^\star_{(c,k+1)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} \vec{x} + \sum\limits_{\vec{x}\in X_1} \vec{x} + \cdots + \sum\limits_{\vec{x}\in X_{k+1}} \vec{x}}{|\mathcal{I}_c| + |X_1| + \cdots + |X_{k+1}|}\\
&= \frac{\vec{\mu}_{(c,k)}N_{(c,k)} + \vec{\mu}_{[X_{k+1}]}|X_{k+1}|}{N_{(c,k)} + |X_{k+1}|}\\
&= \vec{\mu}_{(c,k+1)},
\end{aligned}$$

second, the covariance matrix:

$$\begin{aligned}
\mathbf{\Sigma}^\star_{(c,k+1)} &= \frac{S_1 + S_2}{|\mathcal{I}_c| + |X_1| + \cdots + |X_{k+1}| - 1};\\
S_1 &= \sum_{\vec{x}\in\mathcal{I}_c} (\vec{x} - \vec{\mu}^\star_{(c,k+1)})(\vec{x} - \vec{\mu}^\star_{(c,k+1)})^T\\
&\quad + \sum_{\vec{x}\in X_1} (\vec{x} - \vec{\mu}^\star_{(c,k+1)})(\vec{x} - \vec{\mu}^\star_{(c,k+1)})^T\\
&\quad + \cdots + \sum_{\vec{x}\in X_k} (\vec{x} - \vec{\mu}^\star_{(c,k+1)})(\vec{x} - \vec{\mu}^\star_{(c,k+1)})^T\\
&= (N_{(c,k)} - 1)\mathbf{\Sigma}_{(c,k)} + \Bigg[\frac{N_{(c,k)}(|X_{k+1}|)^2}{(N_{(c,k)} + |X_{k+1}|)^2}\\
&\quad \times (\vec{\mu}_{[X_{k+1}]} - \vec{\mu}_{(c,k)})(\vec{\mu}_{[X_{k+1}]} - \vec{\mu}_{(c,k)})^T\Bigg];\\
S_2 &= \sum_{\vec{x}\in X_{k+1}} (\vec{x} - \vec{\mu}^\star_{(c,k+1)})(\vec{x} - \vec{\mu}^\star_{(c,k+1)})^T\\
&= (|X_{k+1}| - 1)\mathbf{\Sigma}_{[X_{k+1}]} + \Bigg[\frac{|X_{k+1}|(N_{(c,k)})^2}{(N_{(c,k)} + |X_{k+1}|)^2}\\
&\quad \times (\vec{\mu}_{[X_{k+1}]} - \vec{\mu}_{(c,k)})(\vec{\mu}_{[X_{k+1}]} - \vec{\mu}_{(c,k)})^T\Bigg].
\end{aligned}$$

Since S_1 and S_2, then

$$\begin{aligned}
\mathbf{\Sigma}^\star_{(c,k+1)} &= \frac{1}{N_{(c,k)} + |X_{k+1}| - 1}\\
&\quad \times \Big[(N_{(c,k)} - 1)\mathbf{\Sigma}_{(c,k)}\\
&\quad + (|X_{k+1}| - 1)\mathbf{\Sigma}_{[X_{k+1}]} + \boldsymbol{\sigma}_{(c,k+1)}\Big]\\
&= \mathbf{\Sigma}_{(c,k+1)}.
\end{aligned}$$

Since $\vec{\mu}^\star_{(c,k+1)} = \vec{\mu}_{(c,k+1)}$ and $\mathbf{\Sigma}^\star_{(c,k+1)} = \mathbf{\Sigma}_{(c,k+1)}$, then

$$\Theta_{k+1}^\star = \langle(\vec{\mu}_{(c,k+1)}, \mathbf{\Sigma}_{(c,k+1)}) \mid c \in \{1, \ldots, C\}\rangle = \Theta_{k+1},$$

holds for $k + 1$. Hence by the principle mathematical induction $\Theta_t^\star = \Theta_t$ for all $0 \leq t$. □

Proof of Theorem 2. We will prove $\Theta'_t = \Theta_t$ by induction for all $0 \leq t$. Θ'_t is a tuple $\langle(\vec{\mu'}_{(c,t)}, \mathbf{\Sigma'}_{(c,t)}) \mid c \in$

$\{1,\ldots,C\}\rangle$, where:

$$\vec{\mu}'_{(c,t)} = \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c} \vec{x} + \sum\limits_{i\le t}\left(\sum\limits_{\vec{x}\in X_i}\omega_{(c,i)}\vec{x}\right)}{|\mathcal{I}_c| + \sum\limits_{i\le t}|X_i|\omega_{(c,i)}};$$

$$\begin{aligned}\mathbf{\Sigma'}_{(c,t)} = &\frac{1}{|\mathcal{I}_c| + \sum\limits_{i\le t}|X_i|\omega_{(c,i)} - 1}\\ &\times\Bigg[\sum_{\vec{x}\in\mathcal{I}_c}(\vec{x}-\vec{\mu}_c)(\vec{x}-\vec{\mu}_c)^T\\ &+\sum_{i\le t}\left(\sum_{\vec{x}\in X_i}\omega_{(c,i)}(\vec{x}-\vec{\mu}_{[X_i]})(\vec{x}-\vec{\mu}_{[X_i]})^T\right)\Bigg],\end{aligned}$$

and Θ_t is a tuple $\langle(\vec{\mu}_{(c,t)},\mathbf{\Sigma}_{(c,t)}) \mid c\in\{1,\ldots,C\}\rangle$, where:

$$\vec{\mu}_{(c,t)} = \begin{cases}\vec{\mu}_c & \text{if } t=0\\ \frac{N_{(c,t-1)}\vec{\mu}_{(c,t-1)}+|X_t|\omega_{(c,t)}\vec{\mu}_{[X_t]}}{N_{(c,t-1)}+|X_t|\omega_{(c,t)}} & \text{otherwise;}\end{cases}$$

$$\mathbf{\Sigma}_{(c,t)} = \begin{cases}\mathbf{\Sigma}_c & \text{if } t=0\\ \frac{1}{N_{(c,t-1)}+|X_t|\omega_{(c,t)}-1} & \\ \quad\times\Big[(N_{(c,t-1)}-1)\mathbf{\Sigma}_{(c,t-1)} & \text{otherwise.}\\ \quad\quad+\omega_{(c,t)}(|X_t|-1)\mathbf{\Sigma}_{[X_t]}\Big] & \end{cases}$$

Base Case: When $t=0$,

$$\Theta'_0 = \langle(\vec{\mu}'_{(c,0)},\mathbf{\Sigma'}_{(c,0)}) \mid c\in\{1,\ldots,C\}\rangle;$$

$$\vec{\mu}'_{(c,0)} = \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c}\vec{x}}{|\mathcal{I}_c|} = \vec{\mu}_c = \vec{\mu}_{(c,0)};$$

$$\begin{aligned}\mathbf{\Sigma'}_{(c,0)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c}(\vec{x}-\vec{\mu}'_{(c,0)})(\vec{x}-\vec{\mu}'_{(c,0)})^T}{|\mathcal{I}_c|-1}\\ &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c}(\vec{x}-\vec{\mu}_c)(\vec{x}-\vec{\mu}_c)^T}{|\mathcal{I}_c|-1} = \mathbf{\Sigma}_c = \mathbf{\Sigma}_{(c,0)}.\end{aligned}$$

Since $\vec{\mu}'_{(c,0)} = \vec{\mu}_{(c,0)}$ and $\mathbf{\Sigma'}_{(c,0)} = \mathbf{\Sigma}_{(c,0)}$, then

$$\Theta'_0 = \langle(\vec{\mu}_{(c,0)},\mathbf{\Sigma}_{(c,0)}) \mid c\in\{1,\ldots,C\}\rangle = \Theta_0,$$

so the base case holds.

Inductive step: Suppose that $\Theta'_k = \Theta_k$ holds for $0\le k$. We will show that $\Theta'_{k+1} = \Theta_{k+1}$.

$$\Theta'_{k+1} = \langle(\vec{\mu}'_{(c,k+1)},\mathbf{\Sigma'}_{(c,k+1)}) \mid c\in\{1,\ldots,C\}\rangle,$$

starting with the mean vector:

$$\begin{aligned}\vec{\mu}'_{(c,k+1)} &= \frac{\sum\limits_{\vec{x}\in\mathcal{I}_c}\vec{x} + \sum\limits_{\vec{x}\in X_1}\omega_{(c,1)}\vec{x} + \cdots + \sum\limits_{\vec{x}\in X_{k+1}}\omega_{(c,k+1)}\vec{x}}{|\mathcal{I}_c| + |X_1|\omega_{(c,1)} + \cdots + |X_{k+1}|\omega_{(c,k+1)}}\\ &= \frac{\vec{\mu}_{(c,k)}N_{(c,k)} + \vec{\mu}_{[X_{k+1}]}|X_{k+1}|}{N_{(c,k)} + |X_{k+1}|\omega_{(c,k+1)}}\\ &= \vec{\mu}_{(c,k+1)},\end{aligned}$$

second, the covariance matrix:

$$\begin{aligned}\mathbf{\Sigma'}_{(c,k+1)} = &\frac{1}{|\mathcal{I}_c| + |X_1|\omega_{(c,1)} + \cdots + |X_{k+1}|\omega_{(c,k+1)} - 1}\\ &\times\Big[\sum_{\vec{x}\in\mathcal{I}_c}(\vec{x}-\vec{\mu}_c)(\vec{x}-\vec{\mu}_c)^T\\ &+\sum_{\vec{x}\in X_1}\omega_{(c,1)}(\vec{x}-\vec{\mu}_{[X_1]})(\vec{x}-\vec{\mu}_{[X_1]})^T\\ &+\cdots+\sum_{\vec{x}\in X_{k+1}}\omega_{(c,k+1)}(\vec{x}-\vec{\mu}_{[X_{k+1}]})(\vec{x}-\vec{\mu}_{[X_{k+1}]})^T\Big]\\ = &\frac{(N_{(c,k)}-1)\mathbf{\Sigma}_{(c,k)} + \omega_{(c,k+1)}(|X_{k+1}|-1)\mathbf{\Sigma}_{[X_{k+1}]}}{N_{(c,k)} + |X_{k+1}|\omega_{(c,k+1)} - 1}\\ = &\mathbf{\Sigma}_{(c,k+1)}.\end{aligned}$$

Since $\vec{\mu}'_{(c,k+1)} = \vec{\mu}_{(c,k+1)}$ and $\mathbf{\Sigma'}_{(c,k+1)} = \mathbf{\Sigma}_{(c,k+1)}$, then

$$\Theta'_{k+1} = \langle(\vec{\mu}_{(c,k+1)},\mathbf{\Sigma}_{(c,k+1)}) \mid c\in\{1,\ldots,C\}\rangle = \Theta_{k+1},$$

holds for $k+1$. Hence by the principle mathematical induction $\Theta'_t = \Theta_t$ for all $0\le t$ $\square$

B Tables

Dataset	λ						τ					
	LDA			**QDA**			**LDA**			**QDA**		
	FS1	FS2	FS3	FS1	FS2	FS3	FS1	FS2	FS3	FS1	FS2	FS3
NinaPro5	0.6	1.0	1.0	1.0	1.0	1.0	0.4	0.3	0.3	0.0	0.3	0.4
Capgmyo_dbb	1.0	1.0	1.0	1.0	1.0	0.8	0.5	0.5	0.0	0.0	0.5	0.7
MyoArmband	0.7	0.7	1.0	1.0	1.0	1.0	0.7	0.0	0.7	0.5	0.4	0.5
Long-Term 3DC	1.0	1.0	1.0	1.0	1.0	0.01	0.8	0.6	0.6	0.4	0.3	0.3
EMG-EPN-120	1.0	1.0	1.0	1.0	1.0	1.0	0.5	0.4	0.3	0.0	0.3	0.0

Table 4: The best parameters λ and τ of the Nigam-based and thresholding-based classifiers over the five datasets and the three feature sets, using both LDA and QDA.

Dataset	Feature set	LDA			QDA		
		batch classifier using	ours using		batch classifier using	ours using	
		labels	pseudo-labels	labels	labels	pseudo-labels	labels
NinaPro5	FS1	439.06 ± 27.07	0.99±0.15	0.69±0.04	437.53±27.07	0.42±0.19	0.08±0.01
	FS2	524.69 ± 40.06	1.25±0.19	0.90±0.04	523.07±40.04	0.46±0.17	0.09±0.01
	FS3	523.24 ± 43.38	1.23±0.22	0.89±0.05	521.60±43.33	0.44±0.19	0.10±0.01
Capgmyo_dbb	FS1	77.21 ± 3.39	0.79±0.07	0.68±0.06	75.80±3.35	0.19±0.07	0.08±0.01
	FS2	83.66 ± 4.33	0.83±0.09	0.69±0.07	82.23±4.29	0.23±0.09	0.08±0.01
	FS3	86.54 ± 5.84	0.86±0.09	0.71±0.05	85.11±5.80	0.24±0.09	0.09±0.01
MyoArmband	FS1	489.87 ± 21.63	0.78±0.05	0.62±0.02	488.46±21.59	0.17±0.06	0.07±0.01
	FS2	526.97 ± 28.46	0.80±0.04	0.65±0.02	525.52±28.45	0.19±0.08	0.08±0.01
	FS3	532.90 ± 26.68	0.85±0.07	0.67±0.03	531.38±26.66	0.19±0.07	0.08±0.01
Long-Term 3DC	FS1	181.33 ± 8.96	0.90±0.09	0.73±0.05	179.85±8.94	0.24±0.01	0.08±0.01
	FS2	198.33 ± 8.26	0.94±0.09	0.77±0.06	196.81±8.24	0.28±0.10	0.09±0.01
	FS3	198.50 ± 7.87	0.95±0.10	0.79±0.06	196.96±7.85	0.26±0.10	0.09±0.01
EMG-EPN-120	FS1	611.42 ± 200.34	1.88±4.41	1.32±3.62	608.30±199.37	0.48±2.24	0.08±0.01
	FS2	676.62 ± 219.69	1.84±4.23	1.77±4.57	672.41±219.36	0.54±2.33	0.08±0.01
	FS3	662.59 ± 201.74	1.88±4.27	1.09±2.88	659.47±201.23	0.50±2.24	0.17±1.29

Table 5: The average of the updating time [ms] of the batch classifier using labels and our online classifier using labels and pseudo-labels over the five datasets and the three feature sets. Note that all parameter update time were carried out on a desktop computer with an Intel® Core™ i5-8250U processor and 8GB of RAM