A Hinge loss

The multiclass hinge loss is defined by

$$\ell_t(\boldsymbol{W}) = \begin{cases} \max\{1 - m_t(\boldsymbol{W}, y_t), 0\} & \text{if } m_t^\star < \kappa \\ \max\{1 - m_t(\boldsymbol{W}, y_t), 0\} & \text{if } y_t^\star \neq y_t \text{ and } m_t^\star \geq \kappa \\ 0 & \text{if } y_t^\star = y_t \text{ and } m_t^\star \geq \kappa, \end{cases} \tag{5}$$

where $m_t(\boldsymbol{W}, y) = \langle \boldsymbol{W}^y, \boldsymbol{x}_t \rangle - \max_{k \neq y} \langle \boldsymbol{W}^k, \boldsymbol{x}_t \rangle$, $m_t^\star = \max_k m_t(\boldsymbol{W}_t, k)$, and $\kappa \in [0, 1]$. Setting $\kappa = 0$ yields the multiclass hinge loss used in common implementations of the Perceptron. An alternative version of Lemma 1 which holds for the hinge loss can be found in Lemma 3. This can then be used to derive similar results for the hinge loss as for regular surrogate losses.

Lemma 3. *Let* ℓ_t *be the multiclass hinge loss with* $\kappa = \frac{1}{2}$ *and let* $a(\boldsymbol{W}_t, \boldsymbol{x}_t) = \ell\big(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star\big)$*, where* $\ell(\cdot, \boldsymbol{x}_t, y_t) = \ell_t$*. Then* GAPPLETRON *satisfies*

$$\sum_{y \in [K]} p_t'(y) \mathbb{1}[y \neq y_t] \leq \max\left\{\frac{2}{3}, \frac{K-1}{K}\right\} \ell_t(\boldsymbol{W}_t) + \gamma_t.$$

Furthermore, ℓ_t *satisfies* $\|\nabla \ell_t(\boldsymbol{W}_t)\|^2 \leq 4\|\boldsymbol{x}_t\|^2 \ell_t(\boldsymbol{W}_t)$.

Proof of Lemma 3. First, we have that

$$\begin{aligned} \sum_{y \in [K]} p_t'(y) \mathbb{1}[y \neq y_t] &\leq \big(1 - \zeta_t a_t - (1 - \zeta_t)\gamma_t\big) \mathbb{1}[y_t^\star \neq y_t] + \zeta_t a_t \frac{K-1}{K} + (1 - \zeta_t)\gamma_t \\ &\leq (1 - a_t) \mathbb{1}[y_t^\star \neq y_t] + a_t \frac{K-1}{K} + \gamma_t, \end{aligned}$$

where we used that $\zeta \in \{0, 1\}$ and the fact the number of mistakes while uniformly exploring on the dominating set is upper bounded by 1.

To conclude the proof of the first statement we argue that the first two summands of the right hand side are upper bounded by $\frac{K-1}{K} \ell_t(\boldsymbol{W}_t)$. In order to show that, we split the analysis into two cases. In the first case $y_t^\star = y_t$ and the inequality follows by simply substituting $a_t = \ell\big(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star\big) = \ell_t(\boldsymbol{W}_t)$. In the second case $y_t^\star \neq y_t$ and we have that

$$\begin{aligned} m_t^\star + m_t(\boldsymbol{W}_t, y_t) =& \langle \boldsymbol{W}_t^{y_t^\star}, \boldsymbol{x}_t \rangle - \max_{k \neq y_t^\star} \langle \boldsymbol{W}_t^k, \boldsymbol{x}_t \rangle + \langle \boldsymbol{W}_t^{y_t}, \boldsymbol{x}_t \rangle - \max_{k \neq y_t} \langle \boldsymbol{W}_t^k, \boldsymbol{x}_t \rangle \\ =& \langle \boldsymbol{W}_t^{y_t}, \boldsymbol{x}_t \rangle - \max_{k \neq y_t^\star} \langle \boldsymbol{W}_t^k, \boldsymbol{x}_t \rangle \\ \leq& \langle \boldsymbol{W}_t^{y_t}, \boldsymbol{x}_t \rangle - \langle \boldsymbol{W}_t^{y_t}, \boldsymbol{x}_t \rangle = 0 \end{aligned}$$

and thus

$$m_t^\star \leq -m_t(\boldsymbol{W}_t, y_t). \tag{6}$$

Since $y_t^\star \neq y_t$ we also have that

$$\begin{aligned} &(1 - a_t) \mathbb{1}[y_t^\star \neq y_t] + a_t \frac{K-1}{K} \\ &= \Big(1 - \ell\big(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star\big)\Big) + \ell\big(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star\big) \frac{K-1}{K} \\ &= 1 - \frac{1}{K} \ell\big(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star\big) \\ &= 1 - \frac{1}{K} \mathbb{1}[m_t^\star < \kappa]\big(1 - m_t(\boldsymbol{W}_t, y_t^\star)\big). \end{aligned} \tag{7}$$

Now, if $m_t^\star < \kappa$ then by equations (6) and (7) we have

$$(1 - a_t) \mathbb{1}[y_t^\star \neq y_t] + a_t \frac{K-1}{K} = \frac{K-1}{K} + \frac{1}{K} m_t^\star \leq \frac{K-1}{K} (1 + m_t^\star) \leq \frac{K-1}{K} \ell_t(\boldsymbol{W}_t).$$

If $m_t^\star \geq \kappa$, $a_t = 0$. Therefore, by equations (6) and (7) we have that

$$(1 - a_t)\mathbb{1}[y_t^\star \neq y_t] + a_t \frac{K-1}{K} = \frac{1 + m_t^\star}{1 + m_t^\star} \leq \frac{1 - m_t(\boldsymbol{W}_t, y_t)}{1 + \kappa} = \frac{1}{1+\kappa}\ell_t(\boldsymbol{W}_t).$$

Setting $\kappa = \frac{1}{2}$ completes the proof of the first statement.

For the proof of the second statement, first assume that $y_t^\star = y_t$. The case where $m_t^\star \geq \kappa$ is straightforward, so suppose that $m_t^\star < \kappa$, in which case we have that

$$\begin{aligned}\|\nabla \ell_t(\boldsymbol{W}_t)\|^2 \leq& 2\|\boldsymbol{x}_t\|^2 \\ =& \frac{1 - m_t^\star}{1 - m_t^\star}\|\boldsymbol{x}_t\|^2 \\ \leq& 4\|\boldsymbol{x}_t\|^2 \ell_t(\boldsymbol{W}_t).\end{aligned}$$

The case where $y_t^\star \neq y_t$ is evident after observing that $\ell_t(\boldsymbol{W}_t) \geq 1$ in that case. □

B Details of Section 2 (Gappletron)

Lemma 2. *Fix any feedback graph $\mathcal{G}$ and suppose that, for all t, ℓ_t is a regular surrogate loss with respect to ℓ. If $\mathcal{A}$ satisfies equation* (4) *then, for any realization of the randomized predictions $y'_1, \ldots, y'_T$,* GAPPLETRON, *run on $\mathcal{G}$ with gap map a such that $a(\boldsymbol{W}_t, \boldsymbol{x}_t) = \ell(\boldsymbol{W}_t, \boldsymbol{x}_t, y_t^\star)$, satisfies*

$$\begin{aligned}&\sum_{t=1}^T \sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t] \leq \sum_{t=1}^T \widehat{\ell}_t(\boldsymbol{U}) + \sum_{t=1}^T \gamma_t \\ &\quad + \inf_{\eta > 0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^T \left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) - v_t \ell_t(\boldsymbol{W}_t) + \eta v_t^2 L \ell_t(\boldsymbol{W}_t)\right)\right\} \qquad \forall \boldsymbol{U} \in \mathcal{W}\,.\end{aligned}$$

Proof of Lemma 2. By adding and subtracting the surrogate loss of the learner and using the guarantee of $\mathcal{A}$ we have

$$\begin{aligned}&\sum_{t=1}^T \left(\sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t] - \widehat{\ell}_t(\boldsymbol{U})\right) \\ &= \sum_{t=1}^T \left(\sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t] - \widehat{\ell}_t(\boldsymbol{W}_t)\right) + \sum_{t=1}^T \left(\widehat{\ell}_t(\boldsymbol{W}_t) - \widehat{\ell}_t(\boldsymbol{U})\right) \\ &\leq \sum_{t=1}^T \left(\sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t] - \widehat{\ell}_t(\boldsymbol{W}_t)\right) + h(\boldsymbol{U})\sqrt{\sum_{t=1}^T \|\widehat{\boldsymbol{g}}_t\|^2} \\ &\leq \inf_{\eta > 0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^T \left(\sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t] - \widehat{\ell}_t(\boldsymbol{W}_t) + \frac{\eta}{2}\|\widehat{\boldsymbol{g}}_t\|^2\right)\right\},\end{aligned}$$

where in the last inequality we used $\sqrt{ab} = \inf_{\eta>0}\left\{\frac{a}{2\eta} + \frac{\eta}{2}b\right\}$. Recalling that $\widehat{\boldsymbol{g}}_t = v_t \nabla \ell_t(\boldsymbol{W}_t)$, we continue by applying Lemma 1:

$$
\begin{aligned}
&\sum_{t=1}^{T}\left(\sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t] - \widehat{\ell}_t(\boldsymbol{U})\right)\\
&\leq \inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T}\left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t - \widehat{\ell}_t(\boldsymbol{W}_t) + \frac{\eta}{2}\|\widehat{\boldsymbol{g}}_t\|^2\right)\right\}\\
&= \inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T}\left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t - v_t\ell_t(\boldsymbol{W}_t) + \frac{\eta v_t^2}{2}\|\nabla\ell_t(\boldsymbol{W}_t)\|^2\right)\right\}\\
&\leq \inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T}\left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t - v_t\ell_t(\boldsymbol{W}_t) + \eta v_t^2 L\ell_t(\boldsymbol{W}_t)\right)\right\},
\end{aligned}
\tag{8}
$$

where in the final inequality we used equation (3). The lemma's statement follows from rearranging the last inequality. ☐

C Details of Section 3 (Bounds that hold in expectation)

Theorem 4. *Under the conditions of Lemma 2,* GAPPLETRON *with* $\gamma = \frac{1}{2}h(\boldsymbol{U})\sqrt{K\rho L}$ *satisfies:*

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}[y'_t \neq y_t]\right]\\
&\leq \mathbb{E}\left[\sum_{t=1}^{T}\ell_t(\boldsymbol{U})\right] + \max\left\{\frac{2K^2Lh(\boldsymbol{U})^2}{\max\{1,|\mathcal{Q}|\}}, 2\,\mathbb{E}\left[h(\boldsymbol{U})\sqrt{\rho KL\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}\right]\right\}
\end{aligned}
$$

Furthermore, if there exists a $\boldsymbol{U} \in \mathcal{W}$ *such that* $\sum_{t=1}^{T}\ell_t(\boldsymbol{U}) = 0$ *for all realizations of the learners' actions,* GAPPLETRON *with* $\gamma = h(\boldsymbol{U})\sqrt{L\rho}$ *satisfies:*

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}[y'_t \neq y_t]\right]\\
&\leq \mathbb{E}\left[\max\left\{4h(\boldsymbol{U})\sqrt{\rho L\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}, \frac{4KLh(\boldsymbol{U})^2}{\max\{1,|\mathcal{Q}|\}}\right\}\right] - \frac{1}{K}\,\mathbb{E}\left[\sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t)\right].
\end{aligned}
$$

Proof of Theorem 4. Denote by $v_{\max} = \max\{1, \max_t v_t\}$. Observe that $\mathbb{E}_t[v_t] = 1$ and $\mathbb{E}_t[v_t^2] \leq \mathbb{E}_t[v_{\max}]$. We start by applying Lemma 2 and taking expectations:

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}[y'_t \neq y_t]\right]\\
&-\mathbb{E}\left[\sum_{t=1}^{T}\ell_t(\boldsymbol{U})\right] - \mathbb{E}\left[\sum_{t=1}^{T}\gamma_t\right]\\
&\leq \mathbb{E}\left[\inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T}\left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) - v_t\ell_t(\boldsymbol{W}_t) + \eta v_t^2 L\ell_t(\boldsymbol{W}_t)\right)\right\}\right]\\
&\leq \inf_{\eta>0}\left\{\mathbb{E}\left[\frac{h(\boldsymbol{U})^2}{2\eta}\right] + \mathbb{E}\left[\sum_{t=1}^{T}\left(\eta v_{\max}L - \frac{1}{K}\right)\ell_t(\boldsymbol{W}_t)\right]\right\}\\
&\leq \mathbb{E}\left[\frac{v_{\max}KLh(\boldsymbol{U})^2}{2}\right],
\end{aligned}
$$

where the last inequality follows from setting $\eta = \frac{1}{KLv_{\max}}$. By using $\sum_{j=1}^{J} \frac{1}{\sqrt{j}} \le 2\sqrt{J}$ we can see that $\sum_{t=1}^{T} \gamma_t \le 2\gamma\sqrt{\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}$. Now, observe that if $y_t^\star \in \mathcal{Q}$ then $P_t(y_t \in \text{out}(y_t')) \ge \frac{|\mathcal{Q}|}{K}$ and if $y_t^\star \notin \mathcal{Q}$ then $P_t(y_t \in \text{out}(y_t')) \ge \min\left\{\frac{1}{2\rho}, \frac{\gamma_t}{\rho}\right\} \ge \min\left\{\frac{1}{2K}, \frac{\gamma_T}{\rho}\right\}$. This means that

$$v_{\max} \le \max\left\{\frac{\rho}{\gamma_T}, \frac{K}{\max\{1, |\mathcal{Q}|\}}\right\}. \tag{9}$$

Recall that $\gamma_T = \min\{\frac{1}{2}, \gamma/\sqrt{\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}\}$. If $\rho > 1$ then $|\mathcal{Q}| = 0$ which means that if $\frac{1}{2} < \frac{\gamma}{\sqrt{T}}$ then $v_{\max} \le 2K$. On the other hand, if $\rho = 1$ then $|\mathcal{Q}| \ge 1$ which means that if $\frac{1}{2} < \frac{\gamma}{\sqrt{T}}$ then $v_{\max} \le \frac{2K}{|\mathcal{Q}|}$. This in turn means that

$$v_{\max} \le \max\left\{\frac{\rho\sqrt{\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}}{\gamma}, \frac{2K}{\max\{1, |\mathcal{Q}|\}}\right\}. \tag{10}$$

Rearranging the previous inequality and substituting in $\gamma = h(\boldsymbol{U})\sqrt{L|\mathcal{S}|}$,

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[y_t' \ne y_t]\right] &\le \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{U})\right] + \mathbb{E}\left[h(\boldsymbol{U})\sqrt{\rho KL\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}\right] \\
&\quad + \mathbb{E}\left[\max\left\{h(\boldsymbol{U})\sqrt{\rho KL\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}, \frac{2K^2Lh(\boldsymbol{U})^2}{2\max\{1, |\mathcal{Q}|\}}\right\}\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{U})\right] + \mathbb{E}\left[\max\left\{2h(\boldsymbol{U})\sqrt{\rho KL\left|\{t : y_t^\star \notin \mathcal{Q}\}\right|}, \frac{2K^2Lh(\boldsymbol{U})^2}{\max\{1, |\mathcal{Q}|\}}\right\}\right],
\end{aligned}$$

which completes the proof of the first statement of Theorem 4.

Now, in the case where there exists a $\boldsymbol{U} \in \mathcal{W}$ such that $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$ for all realizations of the learners' actions, by the guarantee of $\mathcal{A}$ we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right] &= \mathbb{E}\left[\sum_{t=1}^{T} \left(\widehat{\ell}_t(\boldsymbol{W}_t) - \widehat{\ell}_t(\boldsymbol{U})\right)\right] \\
&\le \inf_{\eta > 0}\left\{\mathbb{E}\left[\frac{h(\boldsymbol{U})^2}{2\eta}\right] + \mathbb{E}\left[\sum_{t=1}^{T} \frac{\eta v_{\max}}{2}\|\nabla \ell_t(\boldsymbol{W}_t)\|^2\right]\right\} \\
&\le \inf_{\eta > 0}\left\{\mathbb{E}\left[\frac{h(\boldsymbol{U})^2}{2\eta}\right] + \mathbb{E}\left[\sum_{t=1}^{T} \eta v_{\max} L \ell_t(\boldsymbol{W}_t)\right]\right\} \\
&\le \mathbb{E}\left[v_{\max} L h(\boldsymbol{U})^2\right] + \tfrac{1}{2}\,\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right],
\end{aligned}$$

where we used that ℓ_t is a regular surrogate loss (in particular equation (3)) and plugged in $\eta = 2(E[v_{\max}]L)^{-1}$. After reordering, the above gives us

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right] \le 2\,\mathbb{E}\left[v_{\max} L h(\boldsymbol{U})^2\right]. \tag{11}$$

Now, by using Lemma 1, (10) and equation (11) we have that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t]\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \frac{K-1}{K} \ell_t(\boldsymbol{W}_t)\right] + \mathbb{E}\left[\sum_{t=1}^{T} \gamma_t\right] - \mathbb{E}\left[\sum_{t=1}^{T} \widehat{\ell}_t(\boldsymbol{W}_t)\right] + \mathbb{E}\left[\sum_{t=1}^{T} \widehat{\ell}_t(\boldsymbol{W}_t)\right] \\
&\leq \mathbb{E}\left[\sum_{t=1}^{T} \gamma_t\right] - \frac{1}{K}\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right] + 2\,\mathbb{E}[v_{\text{max}}] L h(\boldsymbol{U})^2 \\
&\leq 2\,\mathbb{E}\left[h(\boldsymbol{U})\sqrt{L|\mathcal{S}|\{t : y_t^\star \notin \mathcal{Q}\}|}\right] - \frac{1}{K}\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right] \\
&\quad + \mathbb{E}\left[\max\left\{2h(\boldsymbol{U})\sqrt{\rho L|\{t : y_t^\star \notin \mathcal{Q}\}|}, \frac{2KLh(\boldsymbol{U})^2}{\max\{1, |\mathcal{Q}|\}}\right\}\right] \\
&\leq \mathbb{E}\left[\max\left\{4h(\boldsymbol{U})\sqrt{\rho L|\{t : y_t^\star \notin \mathcal{Q}\}|}, \frac{4KLh(\boldsymbol{U})^2}{\max\{1, |\mathcal{Q}|\}}\right\}\right] - \frac{1}{K}\mathbb{E}\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)\right],
\end{aligned}$$

which completes the proof for the second statement of Theorem 4. □

D Details of Section 4 (Bounds that hold with high probability)

We first provide a Lemma due to Beygelzimer et al. (2011) which we use to prove our high-probability bounds.

Lemma 4. *(Beygelzimer et al., 2011, Theorem 1) Let* $Z_1, \ldots, Z_T$ *be a sequence of real-valued random variables. Suppose that* $|Z_t| \leq B$ *and* $\mathbb{E}_t[Z_t] = 0$*. For* $\lambda \in [0, \frac{1}{B}]$ *and* $\delta \in (0, 1)$*, with probability at least* $1 - \delta$*, we have that*

$$\sum_{t=1}^{T} Z_t \leq \lambda(e-2)\sum_{t=1}^{T} \mathbb{E}_t[Z_t^2] + \frac{\ln(1/\delta)}{\lambda}.$$

Theorem 5. *Under the conditions of Lemma 2, with probability at least* $1 - \delta$*,* GAPPLETRON *with* $\gamma = \sqrt{K\rho\left(4\ell_{\text{max}} \ln \frac{1}{\delta'} + Lh(\boldsymbol{U})^2\right)}$ *satisfies:*

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t] \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + 4K \ln \frac{1}{\delta'} \\
&\quad + \max\left\{5\sqrt{K\rho T\left(4\ell_{\text{max}} \ln \frac{1}{\delta'} + Lh(\boldsymbol{U})^2\right)}, \frac{4K^2\left(4\ell_{\text{max}} \ln \frac{1}{\delta'} + Lh(\boldsymbol{U})^2\right)}{\max\{1, |\mathcal{Q}|\}}\right\}
\end{aligned}$$

Furthermore, if there exists a $\boldsymbol{U} \in \mathcal{W}$ *such that* $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$ [4]*, then with probability at least* $1 - \delta$ GAPPLETRON *run with* $\gamma = \sqrt{\rho\left(2Lh(\boldsymbol{U})^2 + 9\ell_{\text{max}} K \ln \frac{1}{2\delta}\right)}$ *satisfies:*

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t] \leq \frac{(K-1)9\ln(1/2\delta)}{2} \\
&\quad + 4\max\left\{\sqrt{\rho\left(2Lh(\boldsymbol{U})^2 + 9\ell_{\text{max}} K \ln \frac{1}{2\delta}\right)T}, \frac{2K\left(2Lh(\boldsymbol{U})^2 + 9\ell_{\text{max}} K \ln \frac{1}{2\delta}\right)}{\max\{1, |\mathcal{Q}|\}}\right\}
\end{aligned}$$

---

[4] Note that $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$, where ℓ_t may depend on the learner's randomness, is a weaker condition than standard separability. For example, if some $\boldsymbol{U}$ satisfies this condition for the standard multiclass hinge loss, then $\boldsymbol{U}$ satisfies the same condition also for our version of the multiclass hinge, see (5).

Proof. Before starting, we find a deterministic upper bound on the right-hand side of (9) in the proof of Theorem 4. First, consider γ_T. By definition it depends on $|\{t : y^*_t \notin \mathcal{Q}\}|$, which is random, however for any realization we can exploit the trivial bound $|\{t : y^*_t \notin \mathcal{Q}\}| \leq T$ to argue that $\gamma_T \geq \min\left\{\frac{1}{2}, \frac{\gamma}{\sqrt{T}}\right\}$. Furthermore, if $\rho > 1$ then $|\mathcal{Q}| = 0$ which means that if $\frac{1}{2} < \frac{\gamma}{\sqrt{T}}$ then $v_{\max} \leq 2K$. On the other hand, if $\rho = 1$ then $|\mathcal{Q}| \geq 1$ which means that if $\frac{1}{2} < \frac{\gamma}{\sqrt{T}}$ then $v_{\max} \leq \frac{2K}{|\mathcal{Q}|}$. With that in mind, we can further bound equation (9):

$$v_{\max} \leq \max\left\{\frac{\rho}{\gamma_T}, \frac{K}{\max\{1, |\mathcal{Q}|\}}\right\} \leq \max\left\{\frac{\rho\sqrt{T}}{\gamma}, \frac{2K}{\max\{1, |\mathcal{Q}|\}}\right\} = V_{\max} \tag{12}$$

As a first step in the actual proof, we study the concentration of the random variables $\mathbb{1}[y'_t \neq y_t]$ around their means $\sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t]$. In order to do so, consider their differences $z_t = \mathbb{1}[y'_t \neq y_t] - \sum_{y \in [K]} p'_t(y)\mathbb{1}[y \neq y_t]$, which have zero mean and are bounded in $[-1, 1]$. By Lemma 1 we have

$$\mathbb{E}_t\left[z_t^2\right] \leq \mathbb{E}_t\left[\mathbb{1}[y'_t \neq y_t]\right] \leq \left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t\right).$$

Thus, we can use Lemma 4 and, with probability at least $1 - \delta'$ we have that for $\eta \in [0, 1]$

$$\begin{aligned}\sum_{t=1}^{T} z_t &\leq \frac{\ln(1/\delta')}{\eta} + \eta \sum_{t=1}^{T}\left(\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t\right) \\ &= \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\left(\frac{\eta'}{K}\ell_t(\boldsymbol{W}_t) + \frac{\eta'}{K-1}\gamma_t\right)\end{aligned} \tag{13}$$

where last equality follows by scaling $\eta = \frac{\eta'}{K-1}$, thus the inequality holds for all $\eta' \in [0, K-1]$. Similarly, we can argue about the concentration of $v_t r_t$ around r_t, where $r_t = \ell_t(\boldsymbol{U}) - \frac{K-\eta'}{K}\ell_t(\boldsymbol{W}_t)$. Note that $\mathbb{E}_t[v_t r_t - r_t] = 0$ and $|v_t r_t - r_t| \leq 2\ell_{\max} V_{\max}$. Moreover

$$\mathbb{E}_t[(v_t r_t - r_t)^2] \leq \mathbb{E}_t\left[(v_t r_t)^2\right] \leq 2V_{\max}\ell_{\max}|r_t| \leq 2V_{\max}\ell_{\max}\left(\ell_t(\boldsymbol{W}_t) + \ell_t(\boldsymbol{U})\right).$$

We can finally apply Lemma 4 on $v_t r_t - r_t$. Therefore, with probability at least $1 - \delta'$, for $\eta \in [0, 1/(2\ell_{\max} V_{\max})]$ it holds that

$$\begin{aligned}\sum_{t=1}^{T}(v_t r_t - r_t) &\leq \frac{\ln(1/\delta)}{\eta} + \eta \sum_{t=1}^{T} 2V_{\max}\ell_{\max}\left(\ell_t(\boldsymbol{W}_t) + \ell_t(\boldsymbol{U})\right) \\ &= \frac{2V_{\max}\ell_{\max}K}{\eta'}\ln(1/\delta) + \frac{\eta'}{K}\sum_{t=1}^{T}\left(\ell_t(\boldsymbol{W}_t) + \ell_t(\boldsymbol{U})\right),\end{aligned} \tag{14}$$

where last inequality is due to scaling $\eta = \frac{\eta'}{2V_{\max}\ell_{\max}K}$, thus the inequality holds for all $\eta' \in [0, K]$.

Choosing $\delta' = \frac{\delta}{2}$, we can conclude that both (13) and (14) hold with probability at least $1-\delta$, for any $\eta \in [0, K-1]$. The rest of the proof consists then in showing that (13) and (14) deterministically imply the claimed bound. In particular, we study two different cases, i.e., when $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) > \sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)$ and its converse.

We first consider $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) \leq \sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)$. By Lemma 2 we find that for any $\eta' \in (0, 1]$

$$\begin{aligned}
&\sum_{t=1}^{T} \sum_{y \in [K]} p'_t(y) \mathbb{1}[y \neq y_t] - \sum_{t=1}^{T} \gamma_t \\
&\quad \leq \sum_{t=1}^{T} v_t \ell_t(\boldsymbol{U}) + \inf_{\eta > 0} \left\{ \frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T} \left(\frac{K-1}{K} \ell_t(\boldsymbol{W}_t) - v_t \ell_t(\boldsymbol{W}_t) + \eta v_t^2 L \ell_t(\boldsymbol{W}_t) \right) \right\} \\
&\quad \leq \sum_{t=1}^{T} v_t \ell_t(\boldsymbol{U}) + \inf_{\eta > 0} \left\{ \frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T} \left(\frac{K-1}{K} \ell_t(\boldsymbol{W}_t) - v_t \ell_t(\boldsymbol{W}_t) + \eta V_{\text{max}} v_t L \ell_t(\boldsymbol{W}_t) \right) \right\} \\
&\quad \leq \sum_{t=1}^{T} v_t \ell_t(\boldsymbol{U}) + \frac{KLV_{\text{max}} h(\boldsymbol{U})^2}{2\eta'} + \sum_{t=1}^{T} \left(\frac{K-1}{K} \ell_t(\boldsymbol{W}_t) - \frac{K-\eta'}{K} v_t \ell_t(\boldsymbol{W}_t) \right) \\
&\quad = \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + \sum_{t=1}^{T} (v_t r_t - r_t) + \frac{KLV_{\text{max}} h(\boldsymbol{U})^2}{2\eta'} + \sum_{t=1}^{T} \frac{\eta' - 1}{K} \ell_t(\boldsymbol{W}_t),
\end{aligned}$$

where we have scaled down $\eta' = KLV_{\text{max}}\eta$. Substituting in (14), we get

$$\begin{aligned}
\sum_{t=1}^{T} \sum_{y \in [K]} p'_t(y) \mathbb{1}[y \neq y_t] \leq{}& \sum_{t=1}^{T} \gamma_t + \left(1 + \frac{\eta'}{K}\right) \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + \frac{KLV_{\text{max}} h(\boldsymbol{U})^2}{2\eta'} \\
&+ \sum_{t=1}^{T} \frac{2\eta' - 1}{K} \ell_t(\boldsymbol{W}_t) + \frac{2V_{\text{max}} \ell_{\text{max}} K}{\eta'} \ln(1/\delta). \qquad (15)
\end{aligned}$$

Equations (13) and (15) are all the ingredients we need to conclude the first case, in fact:

$$\begin{aligned}
&\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] - \ell_t(\boldsymbol{U}) \\
&\leq \sum_{t=1}^{T} \left(\mathbb{1}[y'_t \neq y_t] - \sum_{y \in [K]} p'_t(y) \mathbb{1}[y \neq y_t] \right) + \sum_{t=1}^{T} \sum_{y \in [K]} p'_t(y) \mathbb{1}[y \neq y_t] - \ell_t(\boldsymbol{U}) \\
&\leq \frac{4\eta' - 1}{K} \sum_{t=1}^{T} \ell_t(W_t) + \left(1 + \frac{\eta'}{K-1}\right) \sum_{t=1}^{T} \gamma_t \\
&+ \frac{LKV_{\text{max}} h(\boldsymbol{U})^2}{2\eta'} + \frac{(K-1)\ln(1/\delta')}{\eta'} + \frac{2V_{\text{max}} \ell_{\text{max}} K}{\eta'} \ln(1/\delta) \\
&\leq \frac{5}{4} \sum_{t=1}^{T} \gamma_t + (1 + 8V_{\text{max}} \ell_{\text{max}}) K \ln \frac{1}{\delta'} + 2V_{\text{max}} K \left(L h(\boldsymbol{U})^2 \right),
\end{aligned}$$

where in the last step make the substitution $\eta' = \frac{1}{4}$. Now, we have that $\sum_{t=1}^{T} \gamma_t \leq 2\gamma\sqrt{T}$ and hence we obtain

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] - \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) \\
&\leq 3\gamma\sqrt{T} + K \ln \frac{1}{\delta'} \\
&+ \max\left\{ \frac{\rho K \sqrt{T}}{\gamma}, \frac{2K^2}{\max\{1, |\mathcal{Q}|\}} \right\} \left(8\ell_{\text{max}} \ln \frac{1}{\delta'} + 2Lh(\boldsymbol{U})^2 \right) \\
&\leq K \ln \frac{1}{\delta'} \\
&+ \max\left\{ 5\sqrt{K\rho T \left(4\ell_{\text{max}} \ln \frac{1}{\delta'} + Lh(\boldsymbol{U})^2 \right)}, \frac{4K^2 \left(4\ell_{\text{max}} \ln \frac{1}{\delta'} + Lh(\boldsymbol{U})^2 \right)}{\max\{1, |\mathcal{Q}|\}} \right\}. \qquad (16)
\end{aligned}$$

Consider now the second case, where $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) > \sum_{t=1}^{T} \ell_t(\boldsymbol{W}_t)$. We are still assuming (13) and (14) both hold, even though in this case we need only (13). The $\sum_{t=1}^{T} \sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t]$ term is upper bounded using (13). We have:

$$
\begin{aligned}
&\sum_{t=1}^{T}\big(\mathbb{1}[y'_t \neq y_t] - \ell_t(\boldsymbol{U})\big) \\
&\quad= \sum_{t=1}^{T}(\mathbb{1}[y'_t \neq y_t] - \sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t]) + \sum_{t=1}^{T}\sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t] - \sum_{t=1}^{T}\ell_t(\boldsymbol{U}) \\
&\quad\leq \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\left(\frac{\eta'}{K}\ell_t(\boldsymbol{W}_t) + \frac{\eta'}{K-1}\gamma_t\right) + \sum_{t=1}^{T}\sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t] - \sum_{t=1}^{T}\ell_t(\boldsymbol{U}) \\
&\quad\leq \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\left(\frac{\eta'}{K}\ell_t(\boldsymbol{W}_t) + \frac{\eta'}{K-1}\gamma_t\right) + \sum_{t=1}^{T}\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \gamma_t - \sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t) \\
&\quad= (K-1)4\ln(1/\delta') + \sum_{t=1}^{T}\frac{1}{4K}\ell_t(\boldsymbol{W}_t) + \sum_{t=1}^{T}\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \frac{5}{4}\sum_{t=1}^{T}\gamma_t - \sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t) \\
&\quad\leq 4K\ln(1/\delta') + \frac{5}{4}\sum_{t=1}^{T}\gamma_t \\
&\quad\leq 4K\ln(1/\delta') + 3\gamma\sqrt{T} \\
&\quad= 4K\ln(1/\delta') + 3\sqrt{TK\rho\left(4\ell_{\max}\ln\frac{1}{\delta'} + Lh(\boldsymbol{U})^2\right)}.
\end{aligned}
\tag{17}
$$

The first inequality is due to (13), while the second one to Lemma 1. The third inequality follows from bounding $\frac{1}{4K} \leq \frac{1}{K}$ and finally, the last equality follows by substituting the stated γ.

In order to prove the second statement, we assume there exists a $\boldsymbol{U} \in \mathcal{W}$ such that $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$ for all realizations of the learners' predictions. By the guarantee on $\mathcal{A}$ and inequality (3) we have

$$
\begin{aligned}
\sum_{t=1}^{T}\widehat{\ell}_t(\boldsymbol{W}_t) &= \sum_{t=1}^{T}\left(\widehat{\ell}_t(\boldsymbol{W}_t) - \widehat{\ell}_t(\boldsymbol{U})\right) \\
&\leq \inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \frac{\eta}{2}\sum_{t=1}^{T} v_t^2\|\nabla\ell_t(\boldsymbol{W}_t)\|^2\right\} \\
&\leq \inf_{\eta>0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \eta V_{\max} L\sum_{t=1}^{T} v_t\ell_t(\boldsymbol{W}_t)\right\} \\
&\leq V_{\max} L h(\boldsymbol{U})^2 + \frac{1}{2}\sum_{t=1}^{T} v_t\ell_t(\boldsymbol{W}_t),
\end{aligned}
$$

which, after recalling that $\widehat{\ell}_t(\boldsymbol{W}_t) = v_t\ell_t(\boldsymbol{W}_t)$, and reordering, gives us

$$
\sum_{t=1}^{T}\widehat{\ell}_t(\boldsymbol{W}_t) \leq 2V_{\max} L h(\boldsymbol{U})^2. \tag{18}
$$

By Lemma 4, we have that for $\eta \in [0, \frac{1}{2\ell_{\max}V_{\max}}]$ and $\delta' \in (0,1)$, with probability at least $1-\delta'$

$$
\begin{aligned}
\sum_{t=1}^{T}\left(\ell_t(\boldsymbol{W}_t) - v_t\ell_t(\boldsymbol{W}_t)\right) &\leq \frac{\ln\frac{1}{\delta'}}{\eta} + \eta 2\ell_{\max}V_{\max}\sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t) \\
&= \frac{2\ell_{\max}V_{\max}K\ln\frac{1}{\delta'}}{\eta'} + \frac{\eta'}{K}\sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t),
\end{aligned}
\tag{19}
$$

for all $\eta' \in [0, K]$. By equation (13), with probability at least $1-\delta'$ we have that for all $\eta' \in [0, K-1]$

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] &= \sum_{t=1}^{T} \sum_{y\in[K]} p'_t(y)\mathbb{1}[y\neq y_t] + \sum_{t=1}^{T} \left(\mathbb{1}[y'_t\neq y_t] - \sum_{y\in[K]} p'_t(y)\mathbb{1}[y\neq y_t]\right)\\ &\le \sum_{t=1}^{T}\sum_{y\in[K]} p'_t(y)\mathbb{1}[y\neq y_t] + \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\left(\frac{\eta'}{K}\ell_t(\boldsymbol{W}_t) + \frac{\eta'}{K-1}\gamma_t\right).\end{aligned}$$

We continue by using Lemma 1 to bound $\sum_{y\in[K]} p'_t(y)\mathbb{1}[y \neq y_t]$:

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] \le& \sum_{t=1}^{T} \frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + (1+\frac{\eta'}{K-1})\sum_{t=1}^{T}\gamma_t + \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\frac{\eta'}{K}\ell_t(\boldsymbol{W}_t)\\ =&(1+\frac{\eta'}{K-1})\sum_{t=1}^{T}\gamma_t + \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\frac{\eta'-1}{K}\ell_t(\boldsymbol{W}_t)\\ &+\sum_{t=1}^{T} v_t\ell_t(\boldsymbol{W}_t) + \sum_{t=1}^{T}\left(\ell_t(\boldsymbol{W}_t) - v_t\ell_t(\boldsymbol{W}_t)\right).\end{aligned}$$

By equation (19) and the union bound, with probability at least $1-2\delta'$:

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] \le&(1+\frac{\eta'}{K-1})\sum_{t=1}^{T}\gamma_t + \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\frac{2\eta'-1}{K}\ell_t(\boldsymbol{W}_t)\\ &+\sum_{t=1}^{T} v_t\ell_t(\boldsymbol{W}_t) + \frac{2\ell_{\max}V_{\max}K\ln\frac{1}{\delta'}}{\eta'}\\ \le&(1+\frac{\eta'}{K-1})\sum_{t=1}^{T}\gamma_t + \frac{(K-1)\ln(1/\delta')}{\eta'} + \sum_{t=1}^{T}\frac{2\eta'-1}{K}\ell_t(\boldsymbol{W}_t)\\ &+2V_{\max}Lh(\boldsymbol{U})^2 + \frac{2\ell_{\max}V_{\max}K\ln\frac{1}{\delta'}}{\eta'},\end{aligned}$$

where the final inequality is due to equation (18). We now use $\sum_{t=1}^{T}\gamma_t \le 2\gamma\sqrt{T}$, set $\eta' = \frac{1}{2}$, set $\delta' = \frac{1}{2}\delta$, and use the definition of $V_{\max}$ in equation (12) to continue:

$$\begin{aligned}\sum_{t=1}^{T} \mathbb{1}[y'_t \neq y_t] &\le 3\gamma\sqrt{T} + 2(K-1)\ln(1/2\delta) + 2V_{\max}Lh(\boldsymbol{U})^2 + 4\ell_{\max}V_{\max}K\ln\frac{1}{2\delta}\\ &= 3\gamma\sqrt{T} + (K-1)2\ln(1/2\delta)\\ &\quad+\max\left\{\frac{\rho\sqrt{T}}{\gamma}, \frac{2K}{\max\{1,|\mathcal{Q}|\}}\right\}\left(Lh(\boldsymbol{U})^2 + 4\ell_{\max}K\ln(1/2\delta)\right)\\ &\le (K-1)2\ln(1/2\delta) + 4\max\left\{\sqrt{\rho\left(2Lh(\boldsymbol{U})^2+4\ell_{\max}K\ln\frac{1}{2\delta}\right)T},\right.\\ &\quad\left.\frac{2K\left(2Lh(\boldsymbol{U})^2+4\ell_{\max}K\ln\frac{1}{2\delta}\right)}{\max\{1,|\mathcal{Q}|\}}\right\},\end{aligned}$$

which holds with probability at least $1-\delta$ and completes the proof of the second statement of Theorem 5. □

We now restate Theorem 3, after which we prove it.

Theorem 3. *Under the conditions of Lemma 2, with probability at least* $1-\delta$*,* GAPPLETRON *satisfies*

$$\mathcal{M}_T \le \sum_{t=1}^{T}\ell_t(\boldsymbol{U}) + KLh(\boldsymbol{U})^2 + \frac{3K\ln(1/\delta)}{2} \qquad \forall \boldsymbol{U}\in\mathcal{W}.$$

Furthermore, for all $\boldsymbol{U} \in \mathcal{W}$ *such that* $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$*, then with probability at least* $1 - \delta$*,* GAPPLETRON *satisfies* $\mathcal{M}_T \leq 4Lh(\boldsymbol{U})^2 + \frac{3}{4}\ln\frac{1}{\delta}$*.*

Proof of Theorem 3. Denote by $z_t = \Big(\mathbb{1}[y_t' \neq y_t] - \sum_{y \in [K]} p_t'(y)\mathbb{1}[y \neq y_t]\Big)$. By Lemma 4, for $\lambda \in [0, 1]$, with probability at least $1 - \delta$ we have that

$$\sum_{t=1}^{T} z_t \leq \lambda(e-2)\sum_{t=1}^{T} \mathbb{E}_t\left[\boldsymbol{z}_t^2\right] + \frac{\ln(1/\delta)}{\lambda}$$

Since the variance is bounded by the second moment, we have that

$$\mathbb{E}_t\left[z_t^2\right] \leq \mathbb{E}_t\left[\mathbb{1}[y_t' \neq y_t]\right] \leq \frac{K-1}{K}\ell_t(\boldsymbol{W}_t),$$

where the last inequality is due to Lemma 1. By applying Lemma 2, and recalling that $\gamma_t = 0$ and $v_t = 1$ because we are in the full information setting, we find that with probability at least $1 - \delta$

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t] &\leq \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + \inf_{\eta > 0}\left\{\frac{h(\boldsymbol{U})^2}{2\eta} + \sum_{t=1}^{T}\left(\eta L - \frac{1}{2K}\right)\ell_t(\boldsymbol{W}_t)\right\} \\
&\quad + \left(\lambda\frac{3}{4} - \frac{1}{2K}\right)\sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t) + \frac{\ln(1/\delta)}{\lambda} \\
&\leq \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + KLh(\boldsymbol{U})^2 + \frac{3K\ln(1/\delta)}{2},
\end{aligned}$$

where in the last inequality we used $\lambda = \frac{2}{3K}$, which completes the proof in the non-separable case. In the separable case, when there exists a $\boldsymbol{U} \in \mathcal{W}$ such that $\sum_{t=1}^{T} \ell_t(\boldsymbol{U}) = 0$, we can use Lemma 1 to show that, for some $\lambda \in [0, 1]$, with probability at least $1 - \delta$

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t] &\leq \frac{3\lambda}{4}\sum_{t=1}^{T}\mathbb{E}_t\left[\mathbb{1}[y_t' \neq y_t]\right] + \frac{\ln(1/\delta)}{\lambda} + \sum_{t=1}^{T}\sum_{y \in [K]} p_t'(y)\mathbb{1}[y \neq y_t] \\
&\leq \left(1 + \frac{3\lambda}{4}\right)\sum_{t=1}^{T}\frac{K-1}{K}\ell_t(\boldsymbol{W}_t) + \frac{\ln(1/\delta)}{\lambda} \\
&\leq \frac{3}{4}\ln(1/\delta) + 2\sum_{t=1}^{T}\ell_t(\boldsymbol{W}_t) \\
&\leq \frac{3}{4}\ln(1/\delta) + 4Lh(\boldsymbol{U})^2,
\end{aligned}$$

where in the last inequality we used equation (18) with $V_{\max} = 1$. ☐

E Details of Section 5 (Lower Bounds)

Theorem 6. *In the spam filtering classification setting with smooth hinge loss, the surrogate regret of any (possibly randomized) algorithm satisfies*

$$\mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}[y_t' \neq y_t]\right] = \sum_{t=1}^{T}\ell_t(\widehat{\boldsymbol{U}}) + \Omega\big(\sqrt{T}\big)$$

for some label sequence $y_1, \ldots, y_T \in \{1, 2\}$*, for some sequence of feature vectors* $\boldsymbol{x}_t$ *such that* $\left\|\boldsymbol{x}_t\right\|_2 = \sqrt{2}$ *for all t, and for some* $\widehat{\boldsymbol{U}}$ *such that* $\|\widehat{\boldsymbol{U}}\|_2 \leq \sqrt{5}$*.*

Proof. We adapt an argument from Daniely et al. (2015, Lemma 3). Let $R = \big\lfloor\sqrt{T/2}\big\rfloor$ and divide the T rounds in $2R$ segments $T_1, \ldots, T_{2R}$ of size $T/(2R)$ each (without loss of generality, assume that $2R$ divides T). For each segment T_i, define the components $x_{t,z}$ of the feature vectors $\boldsymbol{x}_t$ at rounds $t \in T_i$ as follows: $x_{t,z} = 1$ if $z \in \{1, i+1\}$ and $x_{t,z} = 0$ otherwise.

Fix an algorithm A and assume $y_t = 1$ for all t. Denote by N_2 the rounds in which A predicts 2. If $\mathbb{E}\left[|N_2|\right] \geq R$, then A makes more than $R = \Omega\big(\sqrt{T}\big)$ mistakes and we are done because

$$\sum_{t=1}^{T} \ell_t(\widehat{\boldsymbol{U}}) = \sum_{t=1}^{T} \max\left\{\left(1 - \hat{U}_1^1 + \hat{U}_1^2\right)^2, 0\right\} = 0$$

for $\widehat{\boldsymbol{U}}$ defined as follows: $\widehat{U}_1^1 = 1$, $\widehat{U}_z^1 = 0$ for $z > 1$, and $\widehat{U}_z^2 = 0$ for all z. Consider then $\mathbb{E}\left[|N_2|\right] \leq R$. Since there are $2R$ segments, we must have that $\mathbb{E}\left[|N_2 \cap T_j|\right] \leq \frac{1}{2}$ for some $j \in [2R]$, because otherwise $\mathbb{E}\left[|N_2|\right] = \mathbb{E}\left[\sum_{i=1}^{2R} |N_2 \cap T_i|\right] > R$. Now, by Markov's inequality we have that $\mathbb{P}(|N_2 \cap T_j| \geq 1) \leq \frac{1}{2}$ which means that A does not predict 2 in segment j with probability at least $\frac{1}{2}$.

Now set $y_t = 2$ for all $t \in T_j$. Since we are in the spam filtering setting, if label 2 is never predicted in segment j, A cannot detect that the label has changed, and so it makes a mistake on each round in T_j, which has length $T/(2R)$. Hence

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t]\right] \geq \frac{T}{2R}\,\mathbb{P}\left(\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t] \geq \frac{T}{2R}\right) \geq \frac{T}{4R} = \Omega\big(\sqrt{T}\big)$$

Define a new comparator $\widehat{\boldsymbol{U}}$ as follows: $\widehat{U}_z^1 = 1$ if $z = 1$ and $\widehat{U}_z^1 = 0$ otherwise, and $\widehat{U}_z^2 = 2$ if $z = j+1$ and $\widehat{U}_z^2 = 0$ otherwise. For the same sequence of labels y_t, we have that

$$\sum_{t=1}^{T} \ell_t(\widehat{\boldsymbol{U}}) = \sum_{t \in T_j} \max\left\{\left(1 - \widehat{U}_1^2 - \widehat{U}_{j+1}^2 + \widehat{U}_1^1 + \widehat{U}_{j+1}^1\right)^2, 0\right\} + \sum_{t \in [T] \setminus T_j} \max\left\{\left(1 - \widehat{U}_1^1 + \widehat{U}_1^2\right)^2, 0\right\}$$

where the sums in the right-hand side are easily seen to be zero. This concludes the proof. ☐

Theorem 7. *Consider the full information setting with smooth hinge loss. For any integer $B \geq 2$, the surrogate regret of any (possibly randomized) algorithm satisfies*

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}[y_t' \neq y_t]\right] \geq \min_{\boldsymbol{U} \in \mathcal{W}} \sum_{t=1}^{T} \ell_t(\boldsymbol{U}) + (B^2 - 1)(K-1) + \frac{K-1}{K}$$

for some label sequence $y_1, \dots, y_T \in [K]$ and for some sequence of feature vectors $\boldsymbol{x}_t$ such that $\left\|\boldsymbol{x}_t\right\|_2 = 1$ for all t, where $\mathcal{W} = \left\{\boldsymbol{W} : \left\|\boldsymbol{W}\right\| \leq B\right\}$.

Proof. We sample the labels y_t uniformly at random for the first $M+1$ rounds, where $M = (B^2 - 1)K^2$. Then we set $y_t = y_{M+1}$ for each $t > M+1$. The feature vectors $\boldsymbol{x}_t$ have $M+1$ components. For $t = 1, \dots, M$ we set the components $x_{t,z}$ of the feature vectors $\boldsymbol{x}_t$ as $x_{t,t} = 1$ and $x_{t,z} = 0$ for $z \neq t$. For each $t \geq M+1$, we set $x_{t,M+1} = 1$ and $x_{t,i} = 0$ for all $i = 1, \dots, M$. We now define a comparator $\widehat{\boldsymbol{U}}$ as follows. For each $z = 1, \dots, M$ we set $\widehat{U}_z^y = \frac{1}{K}$ for $y = y_t$ and $\widehat{U}_z^y = 0$ otherwise. Then we set $\widehat{U}_{M+1}^y = 1$ if $y = y_{M+1}$ and $\widehat{U}_{M+1}^y = 0$ otherwise. Note that, deterministically, $\left\|\widehat{\boldsymbol{U}}\right\|_2^2 = 1 + \sum_{t=1}^{M} \left(U_t^{y_t}\right)^2 = 1 + \frac{M}{K^2} = B^2$. Now fix any (possibly randomized) algorithm A. With these choices, in the first M rounds we have

$$\mathbb{E}\left[\sum_{t=1}^{M} \mathbb{1}[y_t' \neq y_t]\right] - \ell_t(\hat{\boldsymbol{U}}) = M\frac{K-1}{K} - M\left(1 - \frac{1}{K}\right)^2 = M\left(\frac{1}{K} - \frac{1}{K^2}\right).$$

In the next $T - M$ rounds we have

$$\mathbb{E}\left[\sum_{t=M+1}^{T} \mathbb{1}[y_t' \neq y_t]\right] \geq \frac{K-1}{K} \qquad \text{and} \qquad \sum_{t=M+1}^{T} \ell_t(\widehat{\boldsymbol{U}}) = 0\,.$$

The above expectations are both with respect to the random draw of $y_1, \dots, y_{M+1}$ and to A's internal randomization. This implies that there exists a sequence $y_1, \dots, y_{M+1}$ such that the two above bounds hold in expectation with respect to A's internal randomization. Putting the two bounds together concludes the proof. ☐

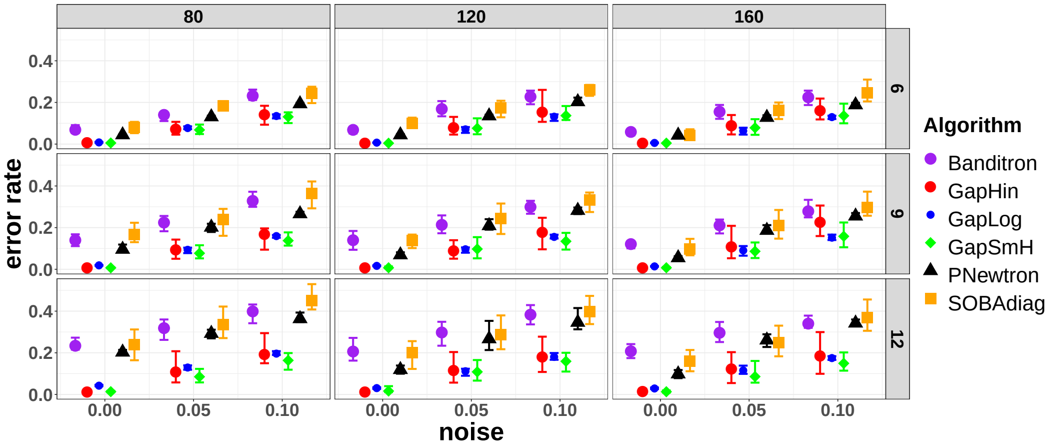

Figure 3: Results of the synthetic experiments for the bandit setting. The parameters of algorithms are set to 1, except for T. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

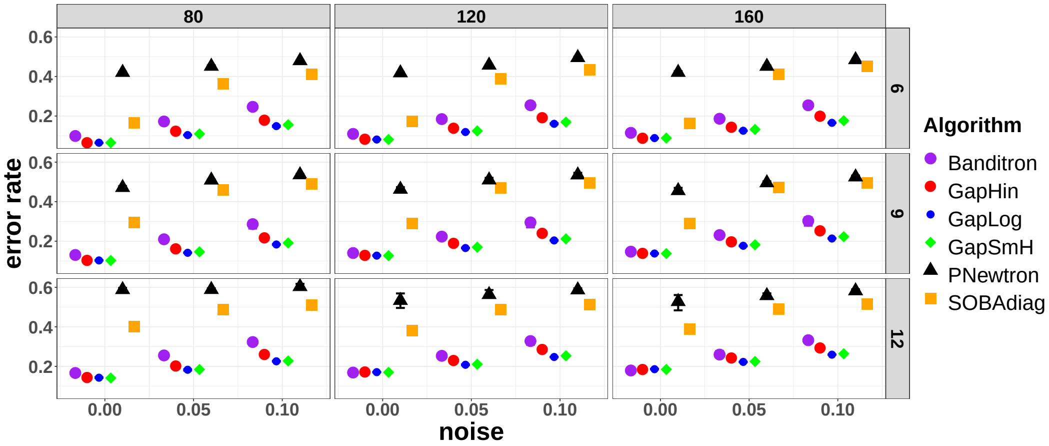

Figure 4: Results of the synthetic experiments for the bandit setting with theoretical tuning. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

F Details of Section 6 (Experiments)

The feature vectors $\boldsymbol{x}_t \in \{0, 1\}^d$ and class labels are generated as follows. For each class we reserve the first $10d'$ bits to generate "keywords". For each class, $1d'$ to $5d'$ of these bits are randomly turned on to represent the keywords for that class. The remaining $30d'$ bits are reserved for unrelated words, of which $5d'$ are randomly turned on. For each t we select a class uniform at random and set $\boldsymbol{x}_t$ to be the feature vector described above. Then, with probability 0, 0.05, or 0.1, we replace the class with a uniformly at random chosen class. We varied between 6, 9, or 12 classes and varied $d' \in [2, 3, 4]$. In the multiclass spam filtering setting we fixed $\mathcal{Q} = \{1\}$, i.e., querying y_t corresponds to predicting label 1.

As suggested by Hazan and Kale (2011), we tuned PNewtron with $\alpha = 10$ and chose the unit ball as domain. For SOBAdiag, we used the adaptive tuning for the exploration rate in the experiment with theoretical tuning and used a fixed exploration rate in the experiment with tuning based only on T.

For the experiments in the partial information setting we tuned the algorithms according to what theory suggests for the worst case. Additionally, we also ran experiments with all parameters set to 1, except for T. Initially we only tuned the algorithms with theoretical tuning, but we found that in the bandit setting two of the algorithms we compare with did not have satisfactory performance.

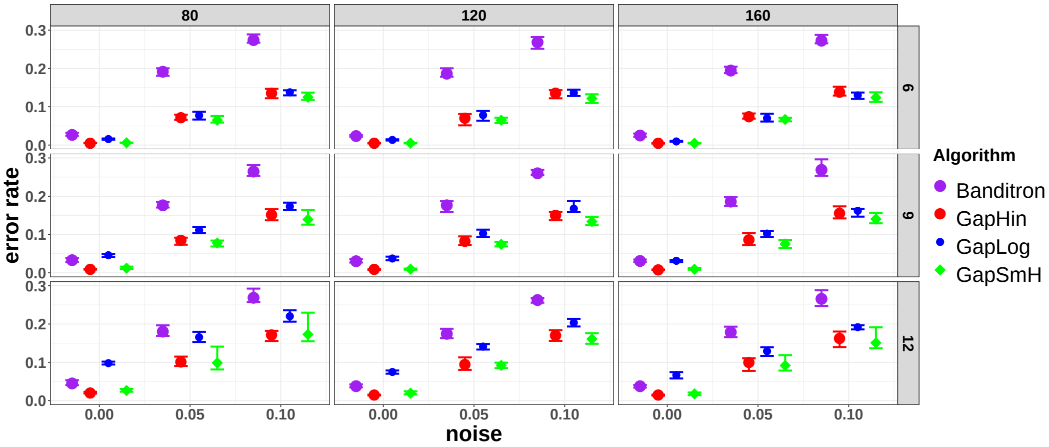

Figure 5: Results of the synthetic experiments for multiclass spam filtering. The parameters of algorithms are set to 1, except for T. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

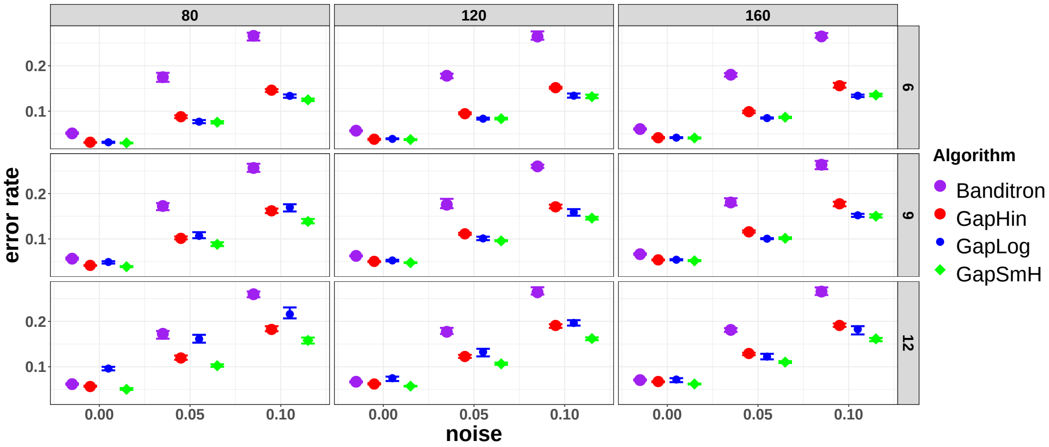

Figure 6: Results of the synthetic experiments for multiclass spam filtering with theoretical tuning. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

All parameters based on (an upper bound on) $\|\boldsymbol{U}\|$ we set to 1 as to not advantage or disadvantage algorithms that did not use tuning based on $\|\boldsymbol{U}\|$. All experiments involving randomness due to the algorithms we repeated ten times.

All experiments were run on a system with 8GB of ram, an Intel i5-6300U CPU, and in python 3.8.5 on a Windows 10 operating system. The results of the experiments are summarized in Figures 3, 4, 5, 6, 7, and 8. We also ran experiments in for the label efficient graph comparing GAPPLETRON with the label efficient PERCEPTRON of Cesa-Bianchi et al. (2006). However, as they assume that labels come without a cost it was not clear how to tune their algorithm. We tried several parameter values which still guarantees sublinear regret in T. However, with none of the choices of parameters the label efficient PERCEPTRON performed as well as GAPPLETRON, so we choose not to report the results of these experiments.

In the experiments with bandit feedback, as we mentioned, Figure 4 shows that with theoretical tuning both PNewtron and SOBAdiag performed poorly. We suspect this is due to the tuning with d, as when we do not tune with d the performance of these algorithms greatly improved (see Figure 3). The error rate of BANDITRON was the lowest with theoretical tuning in roughly half of the experiments. For GapLog and GapSmH the performance also improved when only tuning with T,

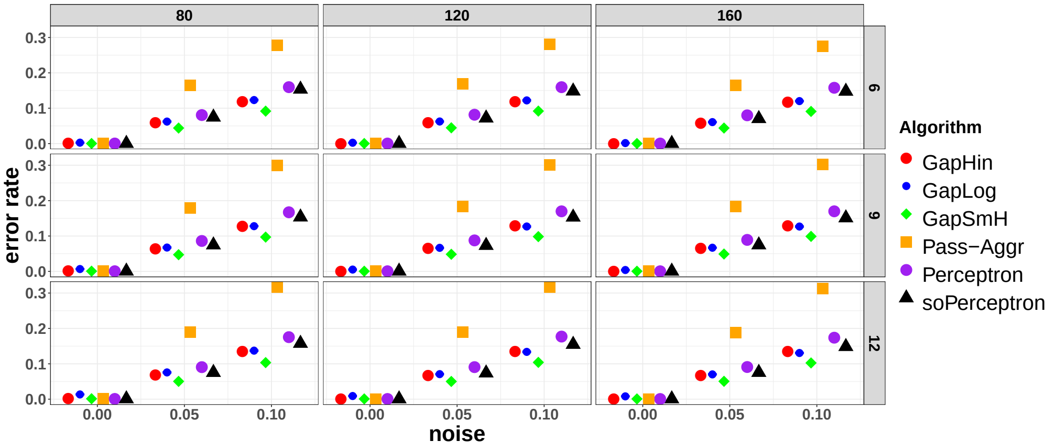

Figure 7: Results of the synthetic experiments for the full information setting. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

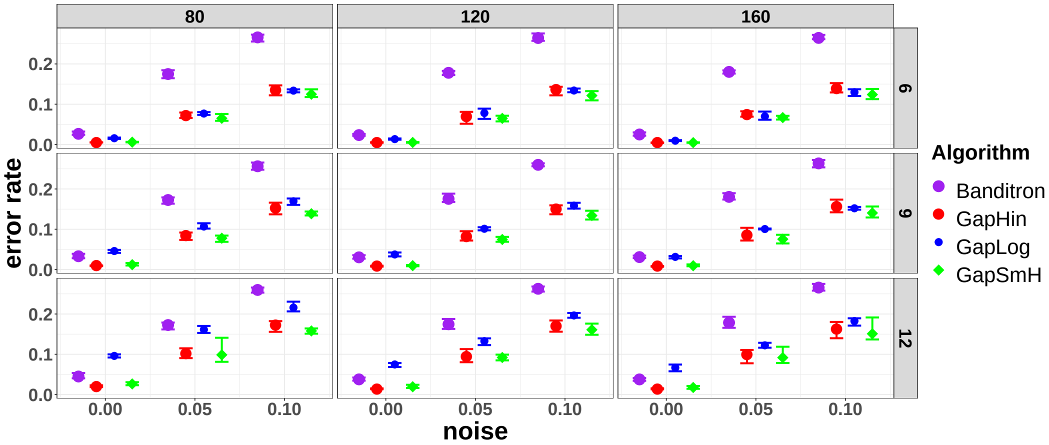

Figure 8: Results of the synthetic experiments for the multiclass spam filtering. The plot shows the best results of algorithms with parameters suggested by theory, or tuned with all parameters set to 1, except for T. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

especially in experiments with no noise. GapHin became more unstable when tuning only with T, as can been seen from the spread of the results. We suspect this is due to the fact that with the hinge loss, GAPPLETRON explores less than with the smooth hinge loss and the logistic loss. Note that with the smooth hinge loss GAPPLETRON explores less than with the logistic loss, which also seems to become apparent from the range of performance of these two versions of GAPPLETRON. With theoretical tuning, in low noise settings Banditron is on par with the performance of all versions of GAPPLETRON, but with high noise GapLog and GapSmH outperform all other algorithms. With tuning that only depends on T, GapLog and GapSmH strictly outperform all other algorithms. Figure 2 contains the results for the best version of each bandit algorithm, which shows that GapLog and GapSmH outperform all other algorithms.

In the multiclass spam filtering setting we compared GAPPLETRON with the importance weighted version of Banditron, which explored the revealing action with probability $\max\{\frac{1}{2}, (X^2/T)^{1/3}\}$ or with probability $\max\{\frac{1}{2}, (1/T)^{1/3}\}$, where the former is the theoretical tuning and $\|\boldsymbol{x}_t\|_2 \leq X$.

In multiclass spam filtering with theoretical tuning (Figure 6), BANDITRON had the lowest error rate in the no-noise experiments. For experiments with noise we see that as K increases the performance

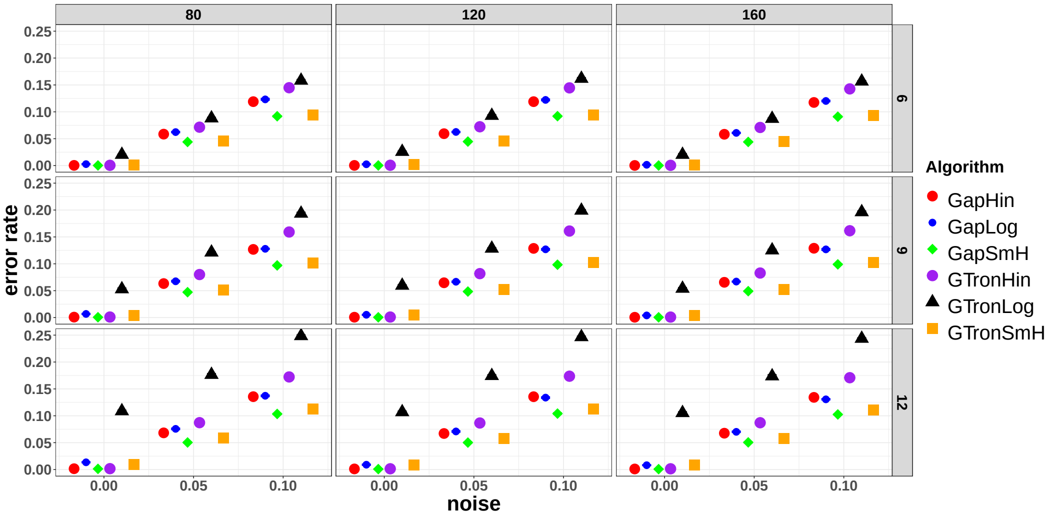

Figure 9: Results of the synthetic experiments for Gaptron and Gappletron in the full information setting. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

of GAPPLETRON deteriorates compared to the performance of BANDITRON. We suspect this is due the $\sqrt{K}$ in the exploration of GAPPLETRON, which does not appear in the exploration of BANDITRON[5]. In Figure 5 we can see that with tuning based solely on T, the spread of the algorithms seems to increase, as was the case in the bandit setting. Either GapHin or GapSmH had the lowest error rate in these experiments, which is also true when comparing the algorithms across the tuning for the exploration rate (Figure 8). The performance of GapLog get worse as K increases, is was the case in the full information setting. We suspect this is due the fact that GapLog explores more than GapHin and GapSmH. While in the bandit setting extra exploration gives additional information, in multiclass spam filtering it does not provide additional information and it only leads to making more mistakes.

In the full information setting we compare GAPPLETRON with the diagonal version of the second-order Perceptron, soPerceptron (Cesa-Bianchi et al., 2005), the multiclass Perceptron, and the passive-aggressive version of the multiclass perceptron (Crammer et al., 2006). In Figure 7, we can see that if there is no label noise, essentially all algorithms find the separating hyperplane. Note that GapLog has the worst performance in this case. This is due to the fact that with the logistic loss, GAPPLETRON never stops with playing at random, leading to sometimes unnecessarily playing the wrong action. We also see this behavior in experiments with label noise, where GapLog performs worse than the other versions of GAPPLETRON, although its performance in still either on par with or better than the non-GAPPLETRON algorithms in these experiments. Overall, in the full information experiments GapSmH appears to have the best performance.

F.1 Gappletron and Gaptron Comparison

We also compared the performance of GAPPLETRON against that of GAPTRON. We use GTronHin as an abbreviation for GAPTRON with the hinge loss, GTronLog as an abbreviation for GAPTRON with the logistic loss, and GTronSmH as an abbreviation for GAPTRON with the smooth hinge loss.

In Figure 9 we see the results of the experiments with full information feedback. For the logistic loss and the hinge loss, GAPPLETRON seems to outperform GAPTRON, especially for larger K. We suspect this is due to the choice of gap map, which is one of the most apparent differences between the two algorithms. The optimizers used by GAPPLETRON and GAPTRON also differ, but this seems to have a smaller impact as for the smooth hinge loss the gap maps coincide and the performances are roughly equivalent.

---

[5] Although no bound exists for this algorithm in literature, one can adapt the proof of Kakade et al. (2008) to prove a $O((X\|\boldsymbol{U}\|)^{1/3}T^{2/3})$ surrogate regret bound.

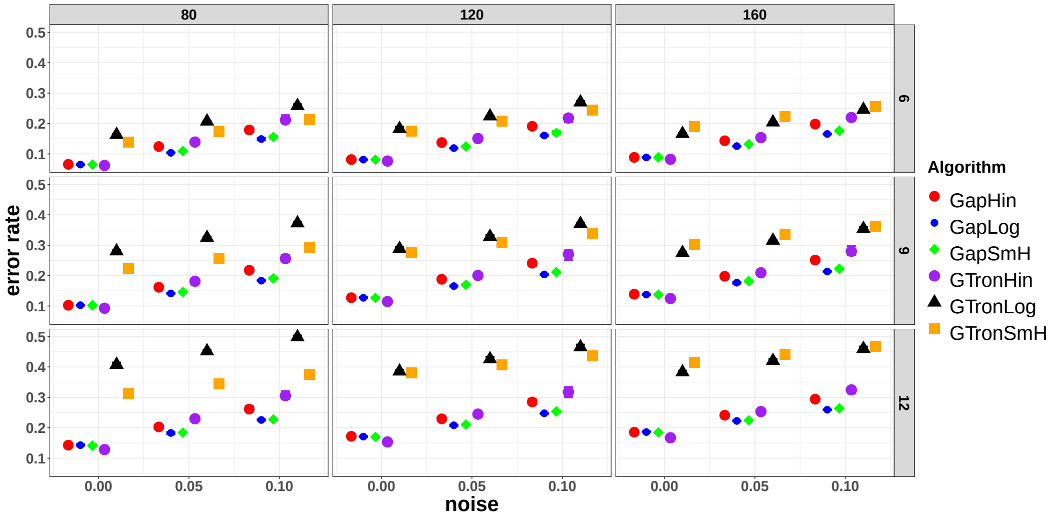

Figure 10: Results of the synthetic experiments for Gaptron and Gappletron with theoretical tuning in the bandit setting. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

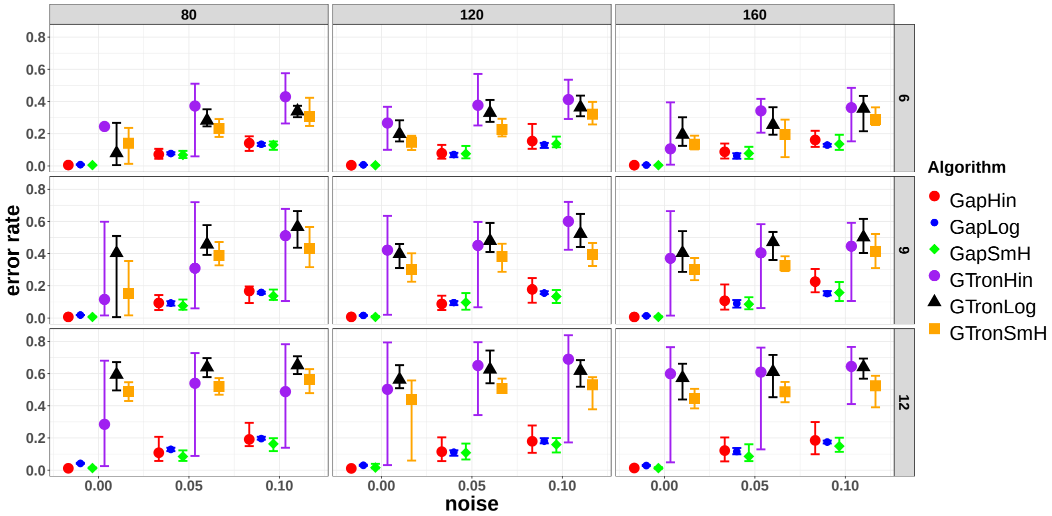

Figure 11: Results of the synthetic experiments for Gaptron and Gappletron in the bandit setting. The parameters of algorithms are set to 1, except for T. The rows are the different values for K and the columns are the different values for d. The whiskers represent the minimum and maximum error rates of the ten repetitions.

In the bandit setting, we did two experiments with GAPTRON: one experiment in which the parameters are set as suggested by theory (Figure 10) and one experiment in which all parameters are set to 1, except for T (Figure 11). With the hinge loss the performance of the algorithms is roughly on par, although with in the separable case GAPTRON slightly outperforms GAPPLETRON whereas in the non-separable setting it is the other way around. For the smooth hinge loss and the logistic loss GAPPLETRON has a smaller error rate, which we suspect is due to the exploration rate. In the experiments where all parameters are set to 1 expect for T GAPPLETRON outperforms GAPTRON.