Appendix

Proof of Lemma 2

Proof. Consider any reward function $w \in \mathbb{R}^K$, any player $i \in N$ and any pair of actions $x, y \in A_i$. We are given that σ matches the expected feature differences of $\tilde{\sigma}$. That is,

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] \tag{16}$$

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] w_k = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] w_k \tag{17}$$

$$\sum_{k=1}^{K} \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] w_k = \sum_{k=1}^{K} \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[f_i^k(y, a_{-i}) - f_i^k(x, a_{-i}) \right] w_k \tag{18}$$

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w \rangle - \langle f_i(x, a_{-i}), w \rangle \right] = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w \rangle - \langle f_i(x, a_{-i}), w \rangle \right] \tag{19}$$

$$\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[u_i(y, a_{-i}) - u_i(x, a_{-i}) \right] = \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[u_i(y, a_{-i}) - u_i(x, a_{-i}) \right] \tag{20}$$

$$\mathbb{E}_{a \sim \sigma} \left[\text{regret}_i(a, \text{switch}_i^{x \to y} | w) \right] = \mathbb{E}_{a \sim \tilde{\sigma}} \left[\text{regret}_i(a, \text{switch}_i^{x \to y} | w) \right] \tag{21}$$

$$\max_{y \in A_i} \mathbb{E}_{a \sim \sigma} \left[\text{regret}_i(a, \text{switch}_i^{x \to y} | w) \right] = \max_{y \in A_i} \mathbb{E}_{a \sim \tilde{\sigma}} \left[\text{regret}_i(a, \text{switch}_i^{x \to y} | w) \right] \tag{22}$$

$$\text{r}_i^{\text{internal}}(x | \sigma, w) = \text{r}_i^{\text{internal}}(x | \tilde{\sigma}, w) \tag{23}$$

$$\sum_{x \in A_i} \text{r}_i^{\text{internal}}(x | \sigma, w) = \sum_{x \in A_i} \text{r}_i^{\text{internal}}(x | \tilde{\sigma}, w) \tag{24}$$

$$\text{R}_i^{\text{swap}}(\sigma | w) = \text{R}_i^{\text{swap}}(\tilde{\sigma} | w) \tag{25}$$

□

Proof of Lemma 3

Proof. For any $i \in N, x \in A_i$ and $w \in \mathcal{R}^k$, let $y' \in A_i$ be an argument that maximizes

$$\max_{y' \in A_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y', a_{-i}), w \rangle - \langle f_i(x, a_{-i}), w \rangle \right]. \tag{26}$$

If $w = 0$, choose $\beta = 1$, otherwise choose β to be the argument that maximizes

$$\max\{\beta : \beta > 0, \forall k, -1 \leq \beta w_k \leq 1\} \tag{27}$$

and let $w' = \beta w$. By construction, w' can be written as a convex combination of the points in $\mathcal{K}_i(x, y' | F, \tilde{\sigma})$. That is,

$$w' = \sum_{v_j \in \mathcal{K}_i(x, y' | F, \tilde{\sigma})} \alpha_j v_j \tag{28}$$

for some $\alpha \geq 0$, $\sum \alpha_j = 1$. Thus, for any $y \in A_i$ we have

$$
\begin{aligned}
&\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w' \rangle - \langle f_i(x, a_{-i}), w' \rangle\right] && (29)\\
&= \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}) - f_i(x, a_{-i}), w' \rangle\right] && (30)\\
&= \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\left\langle f_i(y, a_{-i}) - f_i(x, a_{-i}), \sum_{v_j \in \mathcal{K}_i(x, y' | F, \tilde{\sigma})} \alpha_j v_j \right\rangle\right] && (31)\\
&= \sum_{v_j \in \mathcal{K}_i(x, y' | F, \tilde{\sigma})} \alpha_j \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle (f_i(y, a_{-i}) - f_i(x, a_{-i})), v_j \rangle\right] && (32)\\
&\leq \sum_{v_j \in \mathcal{K}_i(x, y' | F, \tilde{\sigma})} \alpha_j \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle (f_i(y', a_{-i}) - f_i(x, a_{-i})), v_j \rangle\right] && (33)\\
&= \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\left\langle f_i(y', a_{-i}) - f_i(x, a_{-i}), \sum_{v_j \in \mathcal{K}_i(x, y' | F, \tilde{\sigma})} \alpha_j v_j \right\rangle\right] && (34)\\
&= \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y', a_{-i}) - f_i(x, a_{-i}), w' \rangle\right] && (35)\\
&= \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y', a_{-i}), w' \rangle - \langle f_i(x, a_{-i}), w' \rangle\right] && (36)
\end{aligned}
$$

Dividing both sides by β, we get the above inequality in terms of w.

$$
\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w' \rangle - \langle f_i(x, a_{-i}), w' \rangle\right] \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y', a_{-i}), w' \rangle - \langle f_i(x, a_{-i}), w' \rangle\right] \tag{37}
$$

$$
\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\left\langle f_i(y, a_{-i}), \frac{w'}{\beta} \right\rangle - \left\langle f_i(x, a_{-i}), \frac{w'}{\beta} \right\rangle\right] \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\left\langle f_i(y', a_{-i}), \frac{w'}{\beta} \right\rangle - \left\langle f_i(x, a_{-i}), \frac{w'}{\beta} \right\rangle\right] \tag{38}
$$

$$
\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w \rangle - \langle f_i(x, a_{-i}), w \rangle\right] \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y', a_{-i}), w \rangle - \langle f_i(x, a_{-i}), w \rangle\right] \tag{39}
$$

$$
\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[u_i(y, a_{-i} | w) - u_i(x, a_{-i} | w)\right] \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i} | w) \left[u_i(y', a_{-i} | w) - u_i(x, a_{-i} | w)\right] \tag{40}
$$

In particular, this holds for the maximum over $y \in A_i$

$$
\max_{y \in A_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[u_i(y, a_{-i} | w) - u_i(x, a_{-i} | w)\right] \leq \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[u_i(y', a_{-i} | w) - u_i(x, a_{-i} | w)\right] \tag{41}
$$

$$
\mathrm{r}_i^{\text{internal}}(x | \sigma, w) \leq \mathrm{r}_i^{\text{internal}}(x | \tilde{\sigma}, w) \tag{42}
$$

$$
\sum_{x \in A_i} \mathrm{r}_i^{\text{internal}}(x | \sigma, w) \leq \sum_{x \in A_i} \mathrm{r}_i^{\text{internal}}(x | \tilde{\sigma}, w) \tag{43}
$$

$$
\mathrm{R}_i^{\text{swap}}(\sigma | w) \leq \mathrm{R}_i^{\text{swap}}(\tilde{\sigma} | w) \tag{44}
$$

□

Computing w for MaxEntICE-ExpGrad

Given σ, we wish to compute

$$\begin{aligned} &\operatorname*{argmax}_{w} \max_{i \in N} \max_{x,y \in A_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w\rangle - \langle f_i(x, a_{-i}), w\rangle\right] - \\ &\qquad \max_{i \in N} \max_{x,y \in A_i} \sum_{a_{-i} \in \mathcal{A}_{-i}} \tilde{\sigma}(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w\rangle - \langle f_i(x, a_{-i}), w\rangle\right] \\ &\text{subject to:} -1 \leq w \leq 1 \end{aligned} \tag{45}$$

By trying all i, x, y possibilities for the first maximization and all i', x', y' possibilities for the second maximization, we can solve for w using the following series of linear programs.

$$\begin{aligned} \max_{w} &\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w\rangle - \langle f_i(x, a_{-i}), w\rangle\right] - \\ &\sum_{a_{-i'} \in \mathcal{A}_{-i'}} \tilde{\sigma}(x', a_{-i'}) \left[\langle f_{i'}(y', a_{-i'}), w\rangle - \langle f_{i'}(x', a_{-i'}), w\rangle\right] \\ \text{subject to: } &\sum_{a_{-i} \in \mathcal{A}_{-i}} \sigma(x, a_{-i}) \left[\langle f_i(y, a_{-i}), w\rangle - \langle f_i(x, a_{-i}), w\rangle\right] \geq \\ &\sum_{a_{-i''} \in \mathcal{A}_{-i''}} \tilde{\sigma}(x'', a_{-i''}) \left[\langle f_{i''}(y', a_{-i''}), w\rangle - \langle f_{i''}(x', a_{-i''}), w\rangle\right], \text{ and} \\ &\sum_{a_{-i'} \in \mathcal{A}_{-i'}} \tilde{\sigma}(x', a_{-i'}) \left[\langle f_{i'}(y', a_{-i'}), w\rangle - \langle f_{i'}(x', a_{-i'}), w\rangle\right] \geq \\ &\sum_{a_{-i''} \in \mathcal{A}_{-i''}} \tilde{\sigma}(x'', a_{-i''}) \left[\langle f_{i''}(y', a_{-i''}), w\rangle - \langle f_{i''}(x', a_{-i''}), w\rangle\right], \forall i'' \in N, x'', y'' \in A_{i''} \\ &-1 \leq w \leq 1 \end{aligned} \tag{46}$$

The w corresponding to the linear program with the highest optimal value is the w we seek.

Proof of Lemma 4

From [6], we have that Exponentiated Gradient Descent over the d-dimensional simplex has regret no more than $2G_\infty\sqrt{T \log d}$, where G_∞ is a bound on the infinity norm of the gradient with the appropriate choice of η. In our case, $d = A_{\max}^N$. It remains to show that $G_\infty = K$ is an appropriate bound on the gradient. A non-zero entry of the gradient on any iteration has the form

$$\partial\sigma_{x,a_{-i}} = \langle f_i(x, a_{-i}), w\rangle \leq \sum_{k=1}^{K} |f_i(x, a_{-i})| \leq K \tag{47}$$

The analysis of the running time is straightforward. On each iteration, we solve $(NA_{\max}^2)^2$ linear programs with K variables and $2NA_{\max}^2 + 2K$ constraints. Each constraint and computation of the gradient requires $O(A_{\max}^N K)$ work. Updating the weights and the policy requires $O(A^N)$ work. That is, the running time is $O(N^2 A_{\max}^{N+4} K \cdot \text{LP}(K, NA_{\max}^2 + K)$.