A PROOFS

A.1 PROOF OF PROPOSITION 3.3

For probability vectors $\boldsymbol{q}, \boldsymbol{p}, \hat{\boldsymbol{p}} \in \Delta(\mathcal{V})$, define $\mathcal{M}(\boldsymbol{q}, \hat{\boldsymbol{p}}) = \min_{\boldsymbol{p} \in N(\hat{\boldsymbol{p}})} \boldsymbol{q}^\top \log \boldsymbol{p}$, and $\overline{\mathcal{M}}(\hat{\boldsymbol{p}}) = \max_{\boldsymbol{q}} \min_{\boldsymbol{p} \in N(\hat{\boldsymbol{p}})} \boldsymbol{q}^\top \log \boldsymbol{p}$. Then, the t-step total rewards of no-foresight strategy $\mathbb{Q}(\widehat{\mathbb{P}})$ and locally optimal strategy $\widetilde{\mathbb{Q}}(\widehat{\mathbb{P}})$ are respectively given by

$$\mathcal{L}^t(\mathbb{Q}(\widehat{\mathbb{P}}), \mathbb{P}^*(\widehat{\mathbb{P}}, \mathbb{Q})) = \sum_{s=1}^{t} \mathbb{E}_{X_{<s} \sim \mathbb{Q}(\widehat{\mathbb{P}})}[\mathcal{M}(\boldsymbol{q}_s(X_{<s}), \hat{\boldsymbol{p}}_s(X_{<s}))] \coloneqq \mathcal{R}^t(\mathbb{Q}(\widehat{\mathbb{P}}), \widehat{\mathbb{P}}),$$

$$\mathcal{L}^t(\widetilde{\mathbb{Q}}(\widehat{\mathbb{P}}), \mathbb{P}^*(\widehat{\mathbb{P}}, \widetilde{\mathbb{Q}})) = \sum_{s=1}^{t} \mathbb{E}_{X_{<s} \sim \widetilde{\mathbb{Q}}(\widehat{\mathbb{P}})}[\overline{\mathcal{M}}(\hat{\boldsymbol{p}}_s(X_{<s}))] \coloneqq \widetilde{\mathcal{R}}^t(\widehat{\mathbb{P}}).$$

Since $\epsilon < \max_i \hat{p}_i$, $\overline{\mathcal{M}}(\hat{\boldsymbol{p}})$ is always bounded from below. Moreover, as the set-valued mapping $\hat{\boldsymbol{p}} \mapsto N(\hat{\boldsymbol{p}})$ satisfies upper and lower hemicontinuity and $N(\hat{\boldsymbol{p}})$ is compact, $\overline{\mathcal{M}}$ is continuous in $\hat{\boldsymbol{p}}$ by Berge's Maximum Theorem (Aliprantis & Border, 2006), which further implies the continuity of $\widetilde{\mathcal{R}}^t$. Since the space of $\widehat{\mathbb{P}}$ is compact, we conclude that infimum of $\widetilde{\mathcal{R}}^t$ can be attained at some $\widehat{\mathbb{P}}^*$, namely $\inf \widetilde{\mathcal{R}}^t(\widehat{\mathbb{P}}) = \widetilde{\mathcal{R}}^t(\widehat{\mathbb{P}}^*)$.

Now, if $\boldsymbol{q}_t(x_{<t}; \widehat{\mathbb{P}}^*) = \tilde{\boldsymbol{q}}_t(x_{<t}; \widehat{\mathbb{P}}^*)\ \forall t$, we are done. Otherwise, let t_0 be the first step such that $\boldsymbol{q}_{t_0}(x_{<t_0}; \widehat{\mathbb{P}}^*) \neq \tilde{\boldsymbol{q}}_{t_0}(x_{<t_0}; \widehat{\mathbb{P}}^*)$. We have

$$\sum_{s=1}^{t_0-1} \mathbb{E}_{X_{<s} \sim \mathbb{Q}(\widehat{\mathbb{P}}^*)}[\mathcal{M}(\boldsymbol{q}_s(X_{<s}), \hat{\boldsymbol{p}}^*_s(X_{<s}))] = \sum_{s=1}^{t_0-1} \mathbb{E}_{X_{<s} \sim \widetilde{\mathbb{Q}}(\widehat{\mathbb{P}}^*)}[\mathcal{M}(\tilde{\boldsymbol{q}}_s(X_{<s}), \hat{\boldsymbol{p}}^*_s(X_{<s}))],$$
$$\mathbb{E}_{X_{<t_0} \sim \mathbb{Q}(\widehat{\mathbb{P}}^*)}[\mathcal{M}(\boldsymbol{q}_{t_0}(X_{<t_0}), \hat{\boldsymbol{p}}^*_{t_0}(X_{<t_0}))] \leq \mathbb{E}_{X_{<t_0} \sim \widetilde{\mathbb{Q}}(\widehat{\mathbb{P}}^*)}[\mathcal{M}(\tilde{\boldsymbol{q}}_{t_0}(X_{<t_0}), \hat{\boldsymbol{p}}^*_{t_0}(X_{<t_0}))],$$

which implies $\mathcal{R}^{t_0}(\mathbb{Q}(\widehat{\mathbb{P}}^*), \widehat{\mathbb{P}}^*) \leq \widetilde{\mathcal{R}}^{t_0}(\widehat{\mathbb{P}}^*)$. Consider $\widehat{\mathbb{P}}^{**}$ defined as follows. For each $x_{<s} \in \mathcal{V}^{s-1}$,

$$\hat{\boldsymbol{p}}^{**}_s(x_{<s}) = \begin{cases} \hat{\boldsymbol{p}}^*_s(x_{<s}), & s \leq t_0, \\ \hat{\boldsymbol{p}}^*_s(x^*_{<s}) \text{ where } x^*_{<s} = \operatorname{argmin}_{x \in \mathcal{V}^{s-1}} \mathcal{M}(\tilde{\boldsymbol{q}}_s(x), \hat{\boldsymbol{p}}^*_s(x)), & s > t_0. \end{cases}$$

In words, $\widehat{\mathbb{P}}^{**}$ can be understood as shifting the future structure of $\widehat{\mathbb{P}}^*$ after t_0. Since the strategy $\mathbb{Q}(\widehat{\mathbb{P}})$ is defined to have no foresight, we have $\boldsymbol{q}_s(x_{<s}; \widehat{\mathbb{P}}^{**}) = \boldsymbol{q}_s(x_{<s}; \widehat{\mathbb{P}}^*)$ for $s \leq t_0$. Hence,

$$\mathcal{R}^{t_0}(\mathbb{Q}(\widehat{\mathbb{P}}^{**}), \widehat{\mathbb{P}}^{**}) \leq \widetilde{\mathcal{R}}^{t_0}(\widehat{\mathbb{P}}^*) \tag{4}$$

holds as well.

Due to our construction of $\widehat{\mathbb{P}}^{**}$, the future rewards after t_0 satisfy

$$\begin{aligned}\sum_{s=t_0+1}^{T} \mathbb{E}_{X_{<s} \sim \mathbb{Q}(\widehat{\mathbb{P}}^{**})}[\mathcal{M}(\boldsymbol{q}_s(X_{<s}), \hat{\boldsymbol{p}}^{**}_s(X_{<s}))] &\leq \sum_{s=t_0+1}^{T} \max_{x_{<s} \in \mathcal{V}^{s-1}} \mathcal{M}(\boldsymbol{q}_s(x_{<s}), \hat{\boldsymbol{p}}^{**}_s(x_{<s})) \\ &\leq \sum_{s=t_0+1}^{T} \max_{x_{<s} \in \mathcal{V}^{s-1}} \mathcal{M}(\tilde{\boldsymbol{q}}_s(x_{<s}), \hat{\boldsymbol{p}}^{**}_s(x_{<s})) \\ &\leq \sum_{s=t_0+1}^{T} \mathbb{E}_{X_{<s} \sim \widetilde{\mathbb{Q}}(\widehat{\mathbb{P}}^*)}[\mathcal{M}(\tilde{\boldsymbol{q}}_s(X_{<s}), \hat{\boldsymbol{p}}^*_s(X_{<s}))],\end{aligned}$$

namely

$$\mathcal{R}^T(\mathbb{Q}(\widehat{\mathbb{P}}^{**}), \widehat{\mathbb{P}}^{**}) - \mathcal{R}^{t_0}(\mathbb{Q}(\widehat{\mathbb{P}}^{**}), \widehat{\mathbb{P}}^{**}) \leq \widetilde{\mathcal{R}}^T(\widehat{\mathbb{P}}^*) - \widetilde{\mathcal{R}}^{t_0}(\widehat{\mathbb{P}}^*). \tag{5}$$

With (4) and (5), we conclude that

$$\inf_{\widehat{\mathbb{P}}} \mathcal{R}^T(\mathbb{Q}(\widehat{\mathbb{P}}), \widehat{\mathbb{P}}) \leq \mathcal{R}^T(\mathbb{Q}(\widehat{\mathbb{P}}^{**}), \widehat{\mathbb{P}}^{**}) \leq \widetilde{\mathcal{R}}^T(\widehat{\mathbb{P}}^*) = \inf_{\widehat{\mathbb{P}}} \widetilde{\mathcal{R}}^T(\widehat{\mathbb{P}}),$$

which proves the result.

A.2 PROOF OF THEOREM 4.7

We shall only prove the general theorem, as Theorem 4.3 and 4.4 are direct consequences.

Consider the minimization problem

$$\min_{\boldsymbol{p} \in N(\hat{\boldsymbol{p}})} \boldsymbol{q}^\top f(\boldsymbol{p}), \tag{6}$$

where $N(\hat{\boldsymbol{p}}) = \{\boldsymbol{p} \in \Delta(\mathcal{V}) : d_{\text{TV}}(\boldsymbol{p}, \boldsymbol{q}) \leq \epsilon\}$.

The feasible region $N(\hat{\boldsymbol{p}})$ is a convex polytope since it is the intersection of two convex polytopes—the probability simplex $\Delta(\mathcal{V})$ and the ϵ-TV-distance ball $\{\boldsymbol{p} : \frac{1}{2} \|\boldsymbol{p} - \hat{\boldsymbol{p}}\|_1 \leq \epsilon\}$. Moreover, due to concavity of f, it is easy to show that $\boldsymbol{q}^\top f(\boldsymbol{p})$ is concave in $\boldsymbol{p}$. It is well-known that minimizers of a concave function over a polytope are attained at one of the vertices (Horst, 1984). Now, we let $\mathcal{U}$ be the set of the vertices of $N(\hat{\boldsymbol{p}})$.

We will consider the two cases of the theorem separately, due to their differences in the geometry of the feasibility.

Case 1: $\epsilon < \hat{p}_d$, and $\sum_{i=1}^{d-1} \frac{f(\hat{p}_i) - f(\hat{p}_d + \epsilon)}{f(\hat{p}_i) - f(\hat{p}_i - \epsilon)} \geq 1$.

Since $\epsilon < \hat{p}_d$, the set $\mathcal{U}$ can be written as $\mathcal{U} = \{\hat{\boldsymbol{p}} - \epsilon \boldsymbol{e}_i + \epsilon \boldsymbol{e}_j : \ i \neq j\}$. Hence, we have

$$\begin{aligned}
\min_{\boldsymbol{p} \in N(\hat{\boldsymbol{p}})} \boldsymbol{q}^\top f(\boldsymbol{p}) &= \min_{\boldsymbol{p} \in \mathcal{U}} \boldsymbol{q}^\top f(\boldsymbol{p}) \\
&= \boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \min_{i,j : i \neq j} \{q_i \left(f(\hat{p}_i - \epsilon) - f(\hat{p}_i)\right) + q_j \left(f(\hat{p}_j + \epsilon) - f(\hat{p}_j)\right)\} \\
&= \boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) - \max_{i,j : i \neq j} \left\{q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_i)\right\},
\end{aligned}$$

where $g^-(x) \coloneqq f(x) - f(x - \epsilon)$, and $g^+(x) \coloneqq f(x + \epsilon) - f(x)$. Taking this result into our game, the remaining $\boldsymbol{q}$-maximization part is equivalent to

$$\min_{\boldsymbol{q} \in \Delta(\mathcal{V})} \left[-\boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \max_{i,j : i \neq j} \left\{q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_i)\right\} \right]. \tag{7}$$

Ordering of the optimal solution. We claim that any optimal $\boldsymbol{q}^*$ has ordered elements, with $q_1^* \geq \cdots \geq q_d^*$. Observe that both g^+ and g^- are non-increasing, since f is a concave and non-decreasing function. Therefore, if a $\boldsymbol{q}$ has unordered elements, we can rearrange its elements it in descending order, and rearrangement inequality (Hardy et al., 1952) implies that that the term $-\boldsymbol{q}^\top f(\hat{\boldsymbol{p}})$ will decrease. Moreover, by reordering, the term $\max_{i,j:i \neq j} \{q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_i)\}$ will also decrease. This is because

$$\begin{aligned}
\max_{i \neq j} \left\{q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_j)\right\} &= \max_i \left\{ q_i g^-(\hat{p}_i) - \min_{j : j \neq i} q_j g^+(\hat{p}_j) \right\} \\
&= \max_j \left\{ \max_{i : i \neq j} q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_j) \right\},
\end{aligned}$$

Thus, for any fixed i, if we reorder the rest of the elements, $\min_{j \neq i} q_j g^+(\hat{p}_j)$ will increase, making the entire term smaller. Further, by fixing j and reordering by placing q_i in the correct position, $\max_{i \neq j} q_i g^-(\hat{p}_i)$ will decrease. In total, rearranging $\boldsymbol{q}$ in descending order will decrease both terms, resulting in a lower overall objective.

Analyzing KKT optimality. Introducing dual variables $\boldsymbol{\lambda} \in \mathbb{R}_+^d, \nu \in \mathbb{R}$, the Lagrangian of (7) is given by

$$L(\boldsymbol{q}, \boldsymbol{\lambda}, \nu) \coloneqq -\boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \max_{i,j : i \neq j} \left\{q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_j)\right\} - \boldsymbol{\lambda}^\top \boldsymbol{q} + \nu \left(\sum_{i=1}^d q_i - 1 \right).$$

One can check that the objective in (7) is convex in $\boldsymbol{q}$. Moreover, since there exists $\tilde{\boldsymbol{q}} \in \operatorname{relint}(\Delta(\mathcal{V}))$ with $\tilde{\boldsymbol{q}} > 0$, strong duality holds. Therefore, $\boldsymbol{q}^*$ is optimal if and only if there exists $\boldsymbol{\lambda}^*, \nu^*$ such that the following Karush-Kuhn-Tucker (KKT) conditions are satisfied (Boyd & Vandenberghe, 2004):

$$\boldsymbol{0} \in -f(\hat{\boldsymbol{p}}) + \partial \left(\max_{i,j : i \neq j} \left\{q_i^* g^-(\hat{p}_i) - q_j^* g^+(\hat{p}_j)\right\} \right) - \boldsymbol{\lambda}^* + \nu^* \boldsymbol{1}, \qquad \text{(first-order stationarity)}$$

$$\boldsymbol{q}^* \in \Delta(\mathcal{V}), \quad \boldsymbol{\lambda}^* \geq 0, \qquad \text{(primal-dual feasibility)}$$
$$\lambda_i^* q_i^* = 0 \quad \forall i, \qquad \text{(complementary slackness)}$$

where the subdifferential ∂ (Rockafellar, 1970) of the nonsmooth function inside represents the convex hull of the subgradients of the maximizing coordinates, given by

$$\partial\left(\max_{i \neq j}\left\{q_i^* g^-(\hat{p}_i) - q_j^* g^+(\hat{p}_j)\right\}\right) = \operatorname{conv}(\mathcal{D}),$$
$$\mathcal{D} = \left\{g^-(\hat{p}_i)\boldsymbol{e}_i - g^+(\hat{p}_j)\boldsymbol{e}_j : i \neq j,\ q_i^* g^-(\hat{p}_i) - q_j^* g^+(\hat{p}_j) = \max_{i,j: i \neq j}\left\{q_i^* g^-(\hat{p}_i) - q_j^* g^+(\hat{p}_j)\right\}\right\}.$$

Now we show that $\boldsymbol{q}^*$ defined by $q_i^* = \frac{c}{g^-(\hat{p}_i)}\mathbb{1}_{(1 \leq i \leq I^*)}$ satisfies KKT conditions for some dual variables $\boldsymbol{\lambda}^*, \nu^*$, where c is a normalizing constant. Let

$$\mathcal{J} \coloneqq \{i : q_i^* g^-(\hat{p}_i) = c\} = \{1 \leq i \leq I^*\},$$
$$\mathcal{N} \coloneqq \{i : q_i^* g^+(\hat{p}_i) = 0\} = \{I^* < i \leq d\}.$$

Then, as S_I is non-decreasing in I, we have

$$\sum_{k=1}^{I^*-1} \frac{f(\hat{p}_k) - f(\hat{p}_i)}{g^-(\hat{p}_k)} \leq 1, \quad \forall i \in \mathcal{J}, \tag{8}$$

and

$$\sum_{k=1}^{I^*-1} \frac{f(\hat{p}_k) - f(\hat{p}_i)}{g^-(\hat{p}_k)} > 1, \quad \forall i \in \mathcal{N}. \tag{9}$$

Moreover, since

$$S_d = \sum_{k=1}^{d-1} \frac{f(\hat{p}_k) - f(\hat{p}_d)}{g^-(\hat{p}_k)} > \sum_{k=1}^{d-1} \frac{f(\hat{p}_k) - f(\hat{p}_d + \epsilon)}{g^-(\hat{p}_k)} \geq 1,$$

we know that $I^* < d$ must hold, and $\mathcal{N}$ is always non-empty.

To show that KKT conditions are satisfied, it is equivalent to prove that there exist ν^*, $\boldsymbol{\lambda}^* \geq 0$ with $\lambda_i^* = 0$ for $i \in \mathcal{J}$, and coefficients $\gamma_{ij} \geq 0$ for $(i,j) \in \mathcal{J} \times \mathcal{N}$ with $\sum_{i \in \mathcal{J}} \sum_{j \in \mathcal{N}} \gamma_{ij} = 1$ such that

$$-f(\hat{p}_i) + g^-(\hat{p}_i)\left(\sum_{j \in \mathcal{N}} \gamma_{ij}\right)\mathbb{1}_{(i \in \mathcal{J})} - g^+(\hat{p}_i)\left(\sum_{j \in \mathcal{J}} \gamma_{ji}\right)\mathbb{1}_{(i \in \mathcal{N})} - \lambda_i^* \mathbb{1}_{(i \in \mathcal{N})} + \nu^* = 0,$$

which is equivalent to

$$-f(\hat{p}_i) + g^-(\hat{p}_i)\left(\sum_{j \in \mathcal{N}} \gamma_{ij}\right) + \nu^* = 0, \quad i \in \mathcal{J}, \tag{10}$$
$$-f(\hat{p}_i) - g^+(\hat{p}_i)\left(\sum_{j \in \mathcal{J}} \gamma_{ji}\right) + \nu^* = \lambda_i^* \geq 0, \quad i \in \mathcal{N}. \tag{11}$$

The above linear system is satisfied for

$$\nu^* = \left(\sum_{k \in \mathcal{J}} \frac{1}{g^-(\hat{p}_k)}\right)^{-1}\left(\sum_{k \in \mathcal{J}} \frac{f(\hat{p}_k)}{g^-(\hat{p}_k)} - 1\right),$$
$$\gamma_{ij} = \frac{f(\hat{p}_i) - \nu^*}{g^-(\hat{p}_i)}\mathbb{1}_{(j=d)},$$
$$\lambda_i^* = \left(-f(\hat{p}_i) - g^+(\hat{p}_d)\mathbb{1}_{(i=d)} + \nu^*\right)\mathbb{1}_{(i \in \mathcal{N})}.$$

Moreover, (8) and (9) respectively imply that $\gamma_{ij} \geq 0$ and $\lambda_i^* \geq 0$ for all $I^* < i < d$. We also have $\lambda_d^* \geq 0$ because

$$\sum_{k=1}^{d-1} \frac{f(\hat{p}_k) - f(\hat{p}_d) - g^+(\hat{p}_d)}{g^-(\hat{p}_k)} = \sum_{k=1}^{d-1} \frac{f(\hat{p}_k) - f(\hat{p}_d + \epsilon)}{g^-(\hat{p}_k)} \geq 1.$$

Therefore, the above choices of ν^*, γ_{ij}, and λ^* satisfy the linear system and all constraints. Thus, $(\boldsymbol{q}^*, \boldsymbol{\lambda}^*, \nu^*)$ satisfy the KKT conditions, and hence $\boldsymbol{q}^*$ is the optimal solution to problem (f-ODG).

Case 2: $\hat{p}_d \leq \epsilon < \hat{p}_1$, and $\lim_{x \downarrow 0} f(x) = -\infty$.

Let $\mathcal{A} = \{i : \hat{p}_i \leq \epsilon\}$ and $\mathcal{Q} = \{\boldsymbol{q} \in \Delta(\mathcal{V}) : q_i = 0 \ \forall i \in \mathcal{A}\}$. Suppose we use some strategy $\boldsymbol{q} \notin \mathcal{Q}$, i.e., there is some $j \in \mathcal{A}$ such that $q_j \neq 0$. Since $\lim_{x \downarrow 0} f(x) = -\infty$, the adversary can always find $\boldsymbol{p} = \hat{\boldsymbol{p}} - \hat{p}_j \boldsymbol{e}_j$ that makes the objective $-\infty$. Thus, an optimal strategy must come from $\mathcal{Q}$. Similar to Case 1, the $\boldsymbol{p}$-minimization part can be written in terms of the vertex set $\mathcal{U}$ as follows:

$$\begin{aligned}
\min_{\boldsymbol{p} \in N(\hat{\boldsymbol{p}})} \boldsymbol{q}^\top f(\boldsymbol{p}) &= \min_{\boldsymbol{p} \in \mathcal{U}} \boldsymbol{q}^\top f(\boldsymbol{p}) \\
&= \min_{\boldsymbol{p} \in \mathcal{U}_\mathcal{A}} \boldsymbol{q}^\top f(\boldsymbol{p}) \\
&= \boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \min_{(i,j) \in \mathcal{C}} \left\{ q_i \left(f(\hat{p}_i - \epsilon) - f(\hat{p}_i) \right) + q_j \left(f(\hat{p}_j + \epsilon) - f(\hat{p}_j) \right) \right\} \\
&= \boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) - \max_{(i,j) \in \mathcal{C}} \left\{ q_i g^-(\hat{p}_i) - q_j g^+(\hat{p}_i) \right\} \\
&= \boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) - \max_{i \notin \mathcal{A}} q_i g^-(\hat{p}_i), \qquad (12)
\end{aligned}$$

where $\mathcal{U}_\mathcal{A} = \{\hat{\boldsymbol{p}} - \epsilon \boldsymbol{e}_i + \epsilon \boldsymbol{e}_j : i \neq j, i \notin \mathcal{A}\}$, and $\mathcal{C} = \{(i,j) : i \neq j, i \notin \mathcal{A}\}$. (12) follows because $q_j = 0$ for any $j \in \mathcal{A}$. Thus, the problem of interest is equivalent to

$$\min_{\boldsymbol{q} \in \mathcal{Q}} \left[-\boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \max_{i \notin \mathcal{A}} q_i g^-(\hat{p}_i) \right].$$

In other words, we only need to solve $\boldsymbol{q}^*$ from a lower-dimensional problem

$$\min_{\boldsymbol{q} \in \Delta(\mathcal{V}_\mathcal{A})} \left[-\boldsymbol{q}^\top f(\hat{\boldsymbol{p}}) + \max_i q_i g^-(\hat{p}_i) \right],$$

where $\mathcal{V}_\mathcal{A}$ is a truncated vocabulary with $|\mathcal{V}_\mathcal{A}| = d - |\mathcal{A}|$.

Ordering of the optimal solution. Similar to Case 1, an optimal $\boldsymbol{q}^*$ is ordered with $q_1^* \geq \cdots \geq q_d^*$.

Analyzing KKT optimality. The Lagrangian can be similarly defined as

$$L(\boldsymbol{q}, \boldsymbol{\lambda}, \nu) \coloneqq -\boldsymbol{q}^\top f(\hat{p}) + \max_i q_i g^-(\hat{p}_i) - \boldsymbol{\lambda}^\top \boldsymbol{q} + \nu \left(\sum_{i=1}^{d-|\mathcal{A}|} q_i - 1 \right),$$

and strong duality holds as well. The KKT conditions are

$$\begin{aligned}
&\mathbf{0} \in -f(\hat{\boldsymbol{p}}) + \partial \left(\max_i q_i^* g^-(\hat{p}_i) \right) - \boldsymbol{\lambda}^* + \nu^* \mathbf{1}, && \text{(first-order stationarity)} \\
&\boldsymbol{q}^* \in \Delta(\mathcal{V}_\mathcal{A}), \quad \boldsymbol{\lambda}^* \geq 0, && \text{(primal-dual feasibility)} \\
&\lambda_i^* q_i^* = 0 \quad \forall i, && \text{(complementary slackness)}
\end{aligned}$$

where $\partial \left(\max_i q_i^* g^-(\hat{p}_i) \right) \coloneqq \mathrm{conv}\left(\{ g^-(\hat{p}_i) \boldsymbol{e}_i : q_i^* g^-(\hat{p}_i) = \max_i q_i^* g^-(\hat{p}_i) \} \right)$. Let

$$\mathcal{J} = \{i : q_i^* g^-(\hat{p}_i) = c\} = \{1 \leq i \leq I^*\}, \quad \mathcal{N} = \{i : q_i^* g^-(\hat{p}_i) = 0\} = \{I^* < i \leq d - |\mathcal{A}|\},$$

where $c \coloneqq \max_i q_i^* g^-(\hat{p}_i)$. It is sufficient to show that there exist ν^*, $\boldsymbol{\lambda}^* \geq 0$ with $\lambda_i^* = 0$ for $i \in \mathcal{J}$, and coefficients $\gamma_i \geq 0$ for $i \in \mathcal{J}$ with $\sum_{i \in \mathcal{J}} \gamma_i = 1$, such that

$$-f(\hat{p}_i) + \gamma_i g^-(\hat{p}_i) \mathbb{1}_{\{i \in \mathcal{J}\}} - \lambda_i^* \mathbb{1}_{\{i \in \mathcal{N}\}} + \nu^* = 0.$$

This is achieved by setting

$$\nu^* = \left(\sum_{k \in \mathcal{J}} \frac{1}{g^-(\hat{p}_k)} \right)^{-1} \left(\sum_{k \in \mathcal{J}} \frac{f(\hat{p}_k)}{g^-(\hat{p}_k)} - 1 \right),$$

$$\gamma_i = \frac{f(\hat{p}_i) - \nu^*}{g^-(\hat{p}_i)} \geq 0, \quad \text{for } i \in \mathcal{J},$$

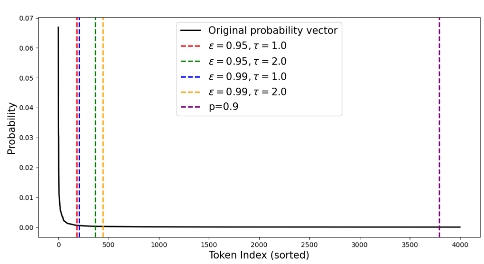
$$\lambda_i^* = (\nu^* - f(\hat{p}_i)) \mathbb{1}_{(i \in \mathcal{N})} \geq 0.$$

Moreover, $\gamma_i \geq 0$ and $\lambda_i^* \geq 0$ follow from the fact that $S_I \leq 1 \ \forall I \in \mathcal{J}$ and $S_I > 1 \ \forall I \in \mathcal{N}$, respectively.

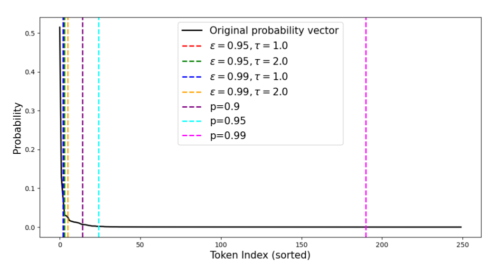
B Additional experiments

In Tables 2 and 3, we present additional experimental results obtained using various choices of ϵ and τ in Game sampling algorithm. These experiments provide further insights into the performance and sensitivity of the model under different parameter settings. We also explored different values of $\epsilon \in \{0.1, 0.3, 0.5, 0.8, 0.9\}$ alongside different τ values. However, since the best performance was consistently achieved with $\epsilon = 0.95$ or $\epsilon = 0.99$, we report only those values here to highlight the effect of changing τ.

As part of this evaluation, we also analyzed the point at which probabilities are truncated and renormalized in Game sampling and Nucleus sampling for a randomly selected article from the WebText test set, using the GPT-2 XL model. The GPT-2 model has a total vocabulary size of 50,000 tokens, so truncating the probability distribution can significantly reduce the set of candidate words for the next token. Figures 1a and 1b illustrate how these sampling strategies truncate the probability distribution. Figure 1a shows the distribution for the next word when using only 1 token as context, along with the index where probabilities are truncated and set to zero. In contrast, Figure 1b presents the distribution for the next word when using the first 35 tokens as context, providing more information for the model to generate the next word. With more context, the model is expected to be more certain about the next word, and the figure highlights the corresponding truncation points. Notably, Game sampling truncates a substantial portion of the 50,000-token distribution and dynamically adjusts the cutoff point based on the shape of the distribution (see Algorithm 1).



(a) context: 1 token



(b) context: 35 tokens

Figure 1: Next-token probability distribution in GPT-2 XL model and truncation threshold of Game sampling and Nucleus sampling.

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	6.874	0.087	0.739
0.95	1.1	7.960	0.058	0.809
0.95	1.5	13.336	0.015	0.898
0.95	2.0	23.592	0.003	0.926
0.95	2.5	40.129	0.002	0.908
0.95	3.0	66.481	0.001	0.815
0.95	3.5	107.544	0.001	0.699
0.95	4.0	172.822	0.001	0.474
0.99	1.0	7.067	0.081	0.746
0.99	1.1	8.275	0.055	0.820
0.99	1.5	14.231	0.012	0.897
0.99	2.0	26.783	0.002	0.917
0.99	2.5	48.508	0.002	0.864
0.99	3.0	89.308	0.001	0.745
0.99	3.5	161.402	0.001	0.529
0.99	4.0	296.453	0.001	0.273

GPT-2 Small

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	6.067	0.048	0.858
0.95	1.1	6.804	0.037	0.883
0.95	1.5	10.423	0.010	0.926
0.95	2.0	17.499	0.003	0.945
0.95	2.5	28.738	0.001	0.919
0.95	3.0	46.973	0.001	0.858
0.95	3.5	78.152	0.001	0.721
0.95	4.0	132.77	0.001	0.475
0.99	1.0	6.176	0.047	0.845
0.99	1.1	6.947	0.033	0.879
0.99	1.5	11.019	0.008	0.941
0.99	2.0	19.482	0.002	0.938
0.99	2.5	34.662	0.002	0.911
0.99	3.0	63.555	0.001	0.792
0.99	3.5	120.889	0	0.497
0.99	4.0	243.844	0	0.257

GPT-2 Medium

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	4.596	0.066	0.823
0.95	1.1	4.972	0.050	0.856
0.95	1.5	6.851	0.013	0.909
0.95	2.0	9.883	0.005	0.942
0.95	2.5	14.084	0.002	0.942
0.95	3.0	19.634	0.002	0.930
0.95	3.5	27.779	0.001	0.913
0.95	4.0	39.256	0.001	0.837
0.99	1.0	4.683	0.066	0.826
0.99	1.1	5.083	0.046	0.861
0.99	1.5	7.130	0.010	0.917
0.99	2.0	10.629	0.006	0.947
0.99	2.5	15.958	0.001	0.947
0.99	3.0	24.128	0.001	0.919
0.99	3.5	37.613	0.001	0.845
0.99	4.0	60.031	0.001	0.685

GPT-2 Large

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	5.146	0.050	0.861
0.95	1.1	5.559	0.033	0.891
0.95	1.5	7.475	0.014	0.935
0.95	2.0	10.541	0.004	0.950
0.95	2.5	14.636	0.002	0.948
0.95	3.0	20.458	0.002	0.929
0.95	3.5	28.410	0.001	0.919
0.95	4.0	39.374	0.001	0.873
0.99	1.0	5.219	0.044	0.852
0.99	1.1	5.660	0.032	0.886
0.99	1.5	7.784	0.010	0.943
0.99	2.0	11.333	0.003	0.958
0.99	2.5	16.690	0.003	0.952
0.99	3.0	24.796	0.002	0.924
0.99	3.5	38.056	0.001	0.885
0.99	4.0	60.236	0.001	0.739

GPT-2 XL

Table 2: Evaluations on the text generated by different types of GPT-2 models using Game sampling under different hyperparameters.

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	5.757	0.069	0.640
0.95	1.1	6.285	0.049	0.670
0.95	1.5	8.528	0.015	0.759
0.95	2.0	12.313	0.005	0.794
0.95	2.5	17.210	0.003	0.811
0.95	3.0	24.362	0.001	0.801
0.95	3.5	33.905	0.002	0.778
0.95	4.0	48.921	0.001	0.664
0.99	1.0	5.897	0.066	0.664
0.99	1.1	6.436	0.046	0.687
0.99	1.5	8.957	0.013	0.762
0.99	2.0	13.263	0.004	0.809
0.99	2.5	19.729	0.002	0.833
0.99	3.0	29.696	0.002	0.791
0.99	3.5	46.506	0	0.720
0.99	4.0	77.289	0.001	0.522

GPT-J-6B

ϵ	τ	Perplexity	Repetition	MAUVE
0.95	1.0	8.500	0.131	0.842
0.95	1.1	9.938	0.134	0.831
0.95	1.5	14.000	0.128	0.858
0.95	2.0	23.875	0.149	0.843
0.95	2.5	36.250	0.162	0.834
0.95	3.0	52.000	0.173	0.813
0.95	3.5	63.750	0.174	0.797
0.95	4.0	87.000	0.182	0.753
0.99	1.0	8.938	0.130	0.831
0.99	1.1	10.250	0.134	0.845
0.99	1.5	15.625	0.136	0.854
0.99	2.0	26.625	0.153	0.840
0.99	2.5	41.750	0.165	0.822
0.99	3.0	60.000	0.181	0.806
0.99	3.5	84.500	0.178	0.759
0.99	4.0	119.500	0.177	0.686

Llama-2-7B

Table 3: Evaluations on the text generated by GPT-J-6B and Llama-2-7B models using Game sampling under different hyperparameters.