A Proof of Admissibility

Proof of Lemma 1. Define functions $\mathbf{Rel}_n : \cup_t [k]^t \to \mathbb{R}$ as

$$\mathbf{Rel}_n(y_1, \dots, y_n) = -\phi(y_1, \dots, y_n)$$

and

$$\mathbf{Rel}_n(y_1, \dots, y_{t-1}) = \mathbb{E}_{y_t \sim \text{Unif}[k]} \mathbf{Rel}_n(y_1, \dots, y_t) + \frac{1}{n}\left(1 - \frac{1}{k}\right), \tag{31}$$

with $\mathbf{Rel}_n(\emptyset)$ being a constant. We desire to prove that there is an algorithm such that

$$\forall \mathbf{y} \in [k]^n, \quad \mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n} \mathbf{1}\{\widehat{y}_t \neq y_t\}\right] - \phi(y_1, \dots, y_n) = 0.$$

Consider the last time step n and write the above expression as

$$\mathbb{E}\left[\frac{1}{n}\sum_{t=1}^{n-1} \mathbf{1}\{\widehat{y}_t \neq y_t\} + \frac{1}{n}\mathbf{1}\{\widehat{y}_n \neq y_n\} + \mathbf{Rel}_n(y_1, \dots, y_n)\right]. \tag{32}$$

Let $\mathbb{E}_{n-1}$ denote the conditional expectation given $\widehat{y}_1, \dots, \widehat{y}_{n-1}$. We shall prove that there exists a randomized strategy for the last step such that for any $y_n \in [k]$,

$$\mathbb{E}_{n-1}\left[\frac{1}{n}\mathbf{1}\{\widehat{y}_n \neq y_n\}\right] + \mathbf{Rel}_n(y_1, \dots, y_n) = \mathbf{Rel}_n(y_1, \dots, y_{n-1}). \tag{33}$$

This last statement is translated as

$$\min_{q_n \in \Delta_k} \max_{y_n \in [k]} \left\{\mathbb{E}_{n-1}\left[\frac{1}{n}\mathbf{1}\{\widehat{y}_n \neq y_n\}\right] + \mathbf{Rel}_n(y_1, \dots, y_n)\right\} = \mathbf{Rel}_n(y_1, \dots, y_{n-1}). \tag{34}$$

Writing $\mathbf{1}\{\widehat{y}_n \neq y_n\} = 1 - \boldsymbol{e}_{\widehat{y}_n}^{\intercal} \boldsymbol{e}_{y_n}$, the left-hand side of (34) is

$$\frac{1}{n} \min_{q_n \in \Delta_k} \max_{y_n \in [k]} \left\{1 - q_n^{\intercal} \boldsymbol{e}_{y_n} + n\mathbf{Rel}_n(y_1, \dots, y_n)\right\}. \tag{35}$$

The stability condition means that we can choose q_n to *equalize* the choices of y_n. Let $\psi(1), \dots, \psi(k)$ be the *sorted* values of

$$n\mathbf{Rel}_n(y_1, \dots, y_{n-1}, 1), \dots, n\mathbf{Rel}_n(y_1, \dots, y_{n-1}, k),$$

in non-increasing order. In view of the stability condition,

$$\sum_{i=1}^{k} (\psi(i) - \psi(k)) \leq 1.$$

Hence, q_n can be chosen so that all $\psi(i) - q_n(i)$ have the same value. One can check that this is the minimizing choice for q_n. Let q_n^* denote this optimal choice. The common value of $\psi(i) - q_n^*(i)$ can then be written as

$$\psi(k) - \frac{1}{k}\left(1 - \sum_{i=1}^{k}(\psi(i) - \psi(k))\right) = \frac{1}{k}\sum_{i=1}^{k}\psi(i) - \frac{1}{k}$$

and hence (35) is equal to

$$\frac{1}{n}\left(1 - \frac{1}{k}\right) + \frac{1}{k}\sum_{i=1}^{k} \mathbf{Rel}_n(y_1, \dots, y_{n-1}, i). \tag{36}$$

This value is precisely $\mathbf{Rel}_n(y_1, \dots, y_{n-1})$, as per Eq. (31), thus verifying (34). Repeating the argument for $t = n - 1$ until $t = 0$, we find that

$$\mathbf{Rel}_n(\emptyset) = -\mathbb{E}\phi + \left(1 - \frac{1}{k}\right) = 0,$$

thus ensuring existence of an algorithm with (32) equal to zero. The other direction of the statement is proved by taking sequences $\mathbf{y}$ uniformly at random from $[k]^n$, concluding the proof. ∎

Proof of Lemma 5. Recall that

$$Y_t = [\nabla_1, \ldots, \nabla_t, \mathbf{0}, \ldots, \mathbf{0}]^\intercal.$$

We can write

$$\begin{aligned}
\mathbf{Rel}_n(\nabla_1, \ldots, \nabla_t) &= \sqrt{\sum_{i=1}^k (Y_t^i)^\intercal M Y_t^i + D^2 \sum_{j=t+1}^n M[j,j]} \\
&= \sqrt{\sum_{i=1}^k (Y_{t-1}^i)^\intercal M Y_{t-1}^i + 2\nabla_t[i] M[t,:] Y_{t-1}^i + \nabla_t^2[i] M[t,t] + D^2 \sum_{j=t+1}^n M[j,j]} \\
&\le \sqrt{\sum_{i=1}^k \left((Y_{t-1}^i)^\intercal M Y_{t-1}^i + 2\nabla_t[i] M[t,:] Y_{t-1}^i\right) + D^2 M[t,t] + D^2 \sum_{j=t+1}^n M[j,j]}
\end{aligned}$$

since $\|\nabla_t\|_2^2 \le D^2$. Hence,

$$\begin{aligned}
&\inf_{\psi_t \in \mathbb{R}^k} \sup_{\|\nabla_t\| \le D} \{\nabla_t^\intercal \psi_t + \mathbf{Rel}_n(\nabla_1, \ldots, \nabla_t)\} \\
&\qquad \le \inf_{\psi_t \in \mathbb{R}^k} \sup_{\|\nabla_t\| \le D} \left\{ \nabla_t^\intercal \psi_t + \sqrt{\sum_{i=1}^k \left((Y_{t-1}^i)^\intercal M Y_{t-1}^i + 2\nabla_t[i] M[t,:] Y_{t-1}^i\right) + D^2 \sum_{j=t}^n M[j,j]} \right\}
\end{aligned}$$

Now the claim is that

$$\psi_t = -\frac{M[t,:] Y_{t-1}}{\sqrt{\sum_{i=1}^k (Y_{t-1}^i)^\intercal M Y_{t-1}^i + D^2 \sum_{j=t}^n M[j,j]}}$$

is the solution to the above minimization problem. To see this, note that for the given ψ_t, the gradient with respect to ∇_t is 0 and hence this ∇_t is the maximizer. Plugging in this solution we get an upper bound on the value

$$\begin{aligned}
&\sup_{\|\nabla_t\| \le D} \left\{ \nabla_t^\intercal \psi_t + \sqrt{\sum_{i=1}^k \left((Y_{t-1}^i)^\intercal M Y_{t-1}^i + 2\nabla_t[i] M[t,:] Y_{t-1}^i\right) + D^2 \sum_{j=t}^n M[j,j]} \right\} \\
&\le \sqrt{\sum_{i=1}^k ((Y_{t-1}^i)^\intercal M Y_{t-1}^i) + D^2 \sum_{j=t}^n M[j,j]} \\
&= \mathbf{Rel}_n(\nabla_1, \ldots, \nabla_{t-1}).
\end{aligned}$$

The bound at the end is given by

$$\mathbf{Rel}_n(\emptyset) = D\sqrt{\text{trace}(M)}.$$

Now once the matrix M is pre-computed, the time complexity per round is $O(t)$. □