A Proof of Proposition 2

We first show the following lemma.

Lemma 1. *Let $\{a_i\}_{i=1}^N$ and $\{b_i\}_{i=1}^N$ be two sequences of real numbers, where $N < \infty$. Let $\{p_i\}_{i=1}^N$ be such that $p_i \geq 0$ for all i and $\sum_{m=1}^N p_m = 1$. Then there exist indices $j, k \in \{1, \cdots, N\}$ (possibly $j = k$) and $p \in [0, 1]$ such that*

$$pa_j + (1-p)a_k \leq \sum_{m=1}^{N} a_m p_m$$

and

$$pb_j + (1-p)b_k \leq \sum_{m=1}^{N} b_m p_m.$$

Proof. We prove the lemma by induction over N. The result is trivial when $N = 1, 2$. Assume that the result holds when $N = n$. In the following we show the case where $N = n + 1$. Let $A = \sum_{m=1}^{n+1} a_m p_m$ and $B = \sum_{m=1}^{n+1} b_m p_m$.

Suppose there exists index $t \in \{1, \cdots, n+1\}$ such that $a_t \leq A$ and $b_t \leq B$, then by choosing $j = k = t$, there is

$$pa_j + (1-p)a_k = a_t \leq A \quad \text{and} \quad pb_j + (1-p)b_k = b_t \leq B.$$

Suppose there exists index $t \in \{1, \cdots, n+1\}$ such that $a_t \geq A$ and $b_t \geq B$. Without loss of generality we can assume $t = n + 1$. If $p_{n+1} = 1$, the result becomes trivial by choosing $j = k = n + 1$. Hence we only consider $p_{n+1} < 1$. Let $p'_i = p_i/(1 - p_{n+1})$ for $i = 1, \cdots, n$, then $\sum_{m=1}^n p'_m = 1$. Applying our assumption to $\{a_i\}_{i=1}^n$, $\{b_i\}_{i=1}^n$ and $\{p'_i\}_{i=1}^n$, we can find $j', k' \in \{1, \cdots, n\}$ and $p' \in [0, 1]$ such that

$$p'a_{j'} + (1-p')a_{k'} \leq \sum_{m=1}^{n} a_m p'_m$$

and

$$p'b_{j'} + (1-p')b_{k'} \leq \sum_{m=1}^{n} b_m p'_m.$$

Notice that

$$\sum_{m=1}^{n} a_m p'_m = \sum_{m=1}^{n} \frac{a_m p_m}{1 - p_{n+1}} \leq \sum_{m=1}^{n+1} a_m p_m = A,$$

and similarly $\sum_{m=1}^n b_m p'_m \leq B$. Therefore by choosing $j = j', k = k'$ and $p = p'$, we arrive at the result.

Consequently, we only have to consider the case where for each $t \in \{1, \cdots, n+1\}$, either $a_t > A, b_t < B$ or $a_t < A, b_t > B$. Without loss of generality, let s be the index such that $a_t > A, b_t < B, \forall t \in \{1, \cdots, s\}$ and $a_t < A, b_t > B, \forall t \in \{s+1, \cdots, n+1\}$. Suppose the result is false, then for any $\ell \in \{1, \cdots, s\}$ and $h \in \{s+1, \cdots, n+1\}$, the following set of inequalities

$$\begin{cases} pa_\ell + (1-p)a_h \leq A \\ pb_\ell + (1-p)b_h \leq B \end{cases}$$

has no solution for p. Since $a_\ell > A > a_h$ and $b_\ell < B < b_h$, this can only happen when

$$\frac{A - a_h}{a_\ell - a_h} < \frac{b_h - B}{b_h - b_\ell}.$$

Rearranging, the above inequality is equivalent to

$$b_h A - b_\ell A + a_\ell B - a_h B + a_h b_\ell - a_\ell b_h < 0. \tag{A1}$$

Let $P' = \sum_{m=1}^{s} p_m$, $A' = \sum_{m=1}^{s} a_m p_m$ and $B' = \sum_{m=1}^{s} b_m p_m$. Multiplying both sides of (A1) by p_ℓ and p_h, and summing over $\ell = 1, \cdots, s$ and $h = s+1, \cdots, n+1$, we have that

$$\begin{aligned} 0 &> \sum_{\ell=1}^{s} \sum_{h=s+1}^{n+1} \Big(b_h p_h p_\ell A - b_\ell p_\ell p_h A + a_\ell p_\ell p_h B - a_h p_h p_\ell B + a_h p_h b_\ell p_\ell - a_\ell p_\ell b_h p_h \Big) \\ &= (B - B')P'A - B'(1 - P')A + A'(1 - P')B - (A - A')P'B + (A - A')B' - A'(B - B') \\ &= 0, \end{aligned}$$

which is a contradiction. Therefore the result holds for $N = n + 1$. $\square$

To show Proposition 2, for each t we construct $\tilde{\theta}_t^*$ that satisfies (i), (ii) and (iii). Notice that, for each $k = 1, \cdots, K$, there is

$$\begin{aligned} \mathbb{E}_{t-1}\Big[\mu(\alpha(\theta_t), \theta^*) \big| \theta_t \in \Theta_k\Big] &= \sum_{\theta \in \Theta_k} \mathbb{P}\big(\theta_t = \theta \big| \theta_t \in \Theta_k\big) \mathbb{E}_{t-1}\Big[\mu\big(\alpha(\theta), \theta^*\big) \big| \theta_t \in \Theta_k\Big] \\ &= \sum_{\theta \in \Theta_k} \mathbb{P}\big(\theta_t = \theta \big| \theta_t \in \Theta_k\big) \mathbb{E}_{t-1}\Big[\mu\big(\alpha(\theta), \theta^*\big)\Big], \end{aligned} \tag{A2}$$

and

$$\begin{aligned} I_{t-1}\left(\psi; Y_{\alpha(\theta_t)} \big| \theta_t \in \Theta_k\right) &= \sum_{\theta \in \Theta_k} \mathbb{P}\big(\theta_t = \theta \big| \theta_t \in \Theta_k\big) I_{t-1}\Big(\psi; Y_{\alpha(\theta)} \big| \theta_t \in \Theta_k\Big) \\ &= \sum_{\theta \in \Theta_k} \mathbb{P}\big(\theta_t = \theta \big| \theta_t \in \Theta_k\big) I_{t-1}\Big(\psi; Y_{\alpha(\theta)}\Big), \end{aligned} \tag{A3}$$

where we used the fact that θ_t is independent of θ^* and ψ.

According to Lemma 1, at stage t, for each $k = 1, \cdots, K$, there exists two parameters $\theta_1^{k,t}, \theta_2^{k,t} \in \Theta_k$ and $r_{k,t} \in [0, 1]$, such that

$$r_{k,t} \cdot \mathbb{E}_{t-1}\left[\mu\big(\alpha(\theta_1^{k,t}), \theta^*\big)\right] + (1 - r_{k,t}) \cdot \mathbb{E}_{t-1}\left[\mu\big(\alpha(\theta_2^{k,t}), \theta^*\big)\right] \le \mathbb{E}_{t-1}\Big[\mu(\alpha(\theta_t), \theta^*) \big| \theta_t \in \Theta_k\Big], \tag{A4}$$

and

$$r_{k,t} \cdot I_{t-1}\left(\psi; Y_{\alpha(\theta_1^{k,t})}\right) + (1 - r_{k,t}) \cdot I_{t-1}\left(\psi; Y_{\alpha(\theta_2^{k,t})}\right) \le I_{t-1}\left(\psi; Y_{\alpha(\theta_t)} \big| \theta_t \in \Theta_k\right). \tag{A5}$$

Let $\tilde{\theta}_t^*$ be a random variable such that

$$\mathbb{P}_{t-1}\Big(\tilde{\theta}_t^* = \theta_1^{k,t} \big| \psi = k\Big) = r_{k,t}, \quad \mathbb{P}_{t-1}\Big(\tilde{\theta}_t^* = \theta_2^{k,t} \big| \psi = k\Big) = 1 - r_{k,t}, \tag{A6}$$

and let $\tilde{\theta}_t$ be an iid copy of $\tilde{\theta}_t^*$. Since the value of $\tilde{\theta}_t^*$ only depends on ψ, (i) is satisfied. Also we have that

$$\begin{aligned} I_{t-1}\Big(\psi; (\tilde{\theta}_t, Y_{\alpha(\tilde{\theta}_t)})\Big) &= I_{t-1}\Big(\psi; \tilde{\theta}_t\Big) + I_{t-1}\Big(\psi; Y_{\alpha(\tilde{\theta}_t)} \big| \tilde{\theta}_t\Big) \\ &\overset{(a)}{=} I_{t-1}\Big(\psi; Y_{\alpha(\tilde{\theta}_t)} \big| \tilde{\theta}_t\Big) \\ &= \sum_{k=1}^{K} \sum_{i=1,2} \mathbb{P}\big(\tilde{\theta}_t = \theta_i^{k,t} \big| \theta_t \in \Theta_k\big) \cdot \mathbb{P}\big(\theta_t \in \Theta_k\big) I_{t-1}\Big(\psi; Y_{\alpha(\theta_i^{k,t})}\Big) \\ &= \sum_{k=1}^{K} \left[r_{k,t} \cdot I_{t-1}\left(\psi; Y_{\alpha(\theta_1^{k,t})}\right) + (1 - r_{k,t}) \cdot I_{t-1}\left(\psi; Y_{\alpha(\theta_2^{k,t})}\right) \right] \cdot \mathbb{P}\big(\theta_t \in \Theta_k\big) \\ &\overset{(b)}{\le} I_{t-1}\left(\psi; Y_{\alpha(\theta_t)} \big| \theta_t \in \Theta_k\right) \cdot \mathbb{P}\big(\theta_t \in \Theta_k\big) \\ &= I_{t-1}\big(\psi; Y_{\alpha(\theta_t)}\big) \\ &\overset{(c)}{=} I_{t-1}\big(\psi; Y_{\alpha(\theta_t)} \big| \theta_t\big) + I_{t-1}\big(\psi; \theta_t\big) = I_{t-1}\big(\psi; (\theta_t, Y_{\alpha(\theta_t)})\big), \end{aligned} \tag{A7}$$

where (a) and (c) follows from that both θ_t and $\tilde{\theta}_t$ are independent of ψ, conditioned on $\tilde{\mathcal{H}}_{t-1}$, and (b) follows from (A5). Therefore (iii) is satisfied.

To show (ii),By construction we have that, at each stage $t = 1, \cdots, T$,

$$D_t = r_{k,t} \cdot \mathbb{E}_{t-1}\left[\mu\big(\alpha(\theta_1^{k,t}), \theta^*\big)\right] + (1 - r_{k,t}) \cdot \mathbb{E}_{t-1}\left[\mu\big(\alpha(\theta_2^{k,t}), \theta^*\big)\right] - \mathbb{E}_{t-1}\Big[\mu(\alpha(\theta_t), \theta^*)\big|\theta_t \in \Theta_k\Big] \leq 0.$$

Hence there is

$$\begin{aligned}
\mathbb{E}_{t-1}\Big[R(Y_{\alpha(\tilde{\theta}_t)}) - R(Y_{\alpha(\theta_t)})\Big] &= \mathbb{E}_{t-1}\Big[\mu\big(\alpha(\tilde{\theta}_t), \theta^*\big) - \mu\big(\alpha(\theta_t), \theta^*\big)\Big] \\
&= \sum_{k=1}^{K} \mathbb{P}\Big(\theta_t \in \Theta_k\Big) \cdot \mathbb{E}_{t-1}\Big[\mu\big(\alpha(\tilde{\theta}_t), \theta^*\big) - \mu\big(\alpha(\theta_t), \theta^*\big) \mid \theta_t \in \Theta_k\Big] \\
&= \sum_{k=1}^{K} \mathbb{P}\Big(\theta_t \in \Theta_k\Big) \cdot D_t \leq 0. \qquad \text{(A8)}
\end{aligned}$$

Therefore we arrive at

$$\begin{aligned}
& \mathbb{E}_{t-1}\Big[R^* - R(Y_{\alpha(\theta_t)})\Big] - \mathbb{E}_{t-1}\Big[R(Y_{\alpha(\tilde{\theta}_t^*)}) - R(Y_{\alpha(\tilde{\theta}_t)})\Big] \\
= \; & \mathbb{E}_{t-1}\Big[R(Y_{\alpha(\theta^*)}) - R(Y_{\alpha(\tilde{\theta}_t^*)})\Big] + \mathbb{E}_{t-1}\Big[R(Y_{\alpha(\tilde{\theta}_t)}) - R(Y_{\alpha(\theta_t)})\Big] \\
\leq \; & \mathbb{E}_{t-1}\Big[\mu(\alpha(\theta^*), \theta^*) - \mu(\alpha(\tilde{\theta}_t^*), \theta^*)\Big] \\
\leq \; & \epsilon, \qquad \text{(A9)}
\end{aligned}$$

where the final step comes from the fact that θ^* and $\tilde{\theta}_t^*$ are always in the same partition.

B Proof of Proposition 3

First, for two random parameters θ and θ' we define

$$\tilde{\Gamma}_t(\theta; \theta') = \frac{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta)}) - R(Y_{\alpha(\theta')})\big]^2}{I_{t-1}\big(\theta; (\theta', Y_{\alpha(\theta')})\big)}, \qquad \text{(A10)}$$

where the subscript $t-1$ indicates the corresponding value under base measure $\tilde{\mathcal{H}}_{t-1}$. From the definition, $\tilde{\Gamma}_t(\theta; \theta')$ is a random variable measurable with respect to $\sigma(\tilde{\mathcal{H}}_{t-1})$.

Lemma 2. *We have that, for each $t = 1, \cdots, T$,*

$$\begin{aligned}
I_{t-1}\Big(\tilde{\theta}_t^*; (\tilde{\theta}_t, Y_{\alpha(\tilde{\theta}_t)})\Big) &= \sum_{i=1}^{m} \mathbb{P}_{t-1}\Big(\tilde{\theta}_t = \theta^i\Big) I_{t-1}\Big(\tilde{\theta}_t^*; Y_{\alpha(\theta^i)}\Big) \\
&\geq 2 \sum_{i=1}^{m} \sum_{j=1}^{m} \mathbb{P}_{t-1}\Big(\tilde{\theta}_t^* = \theta^i\Big) \mathbb{P}_{t-1}\Big(\tilde{\theta}_t^* = \theta^j\Big) \cdot \\
&\quad \Big\{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})|\tilde{\theta}_t^* = \theta^j\big] - \mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big]\Big\}
\end{aligned}$$

and

$$\mathbb{E}_{t-1}\big[R(Y_{\tilde{\theta}_t^*}) - R(Y_{\tilde{\theta}_t})\big] = \sum_{i=1}^{m} \mathbb{P}_{t-1}\Big(\tilde{\theta}_t^* = \theta^i\Big) \Big\{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})|\tilde{\theta}_t^* = \theta^i\big] - \mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big]\Big\},$$

almost surely.

Proof. For each $t = 1, \cdots, T$, there is

$$\begin{aligned}
I_{t-1}\Big(\tilde\theta_t^*; (\tilde\theta_t, Y_{\alpha(\tilde\theta_t)})\Big) &= I_{t-1}\Big(\tilde\theta_t^*; \tilde\theta_t\Big) + I_{t-1}\Big(\tilde\theta_t^*; Y_{\alpha(\tilde\theta_t)}|\tilde\theta_t\Big) \\
&\stackrel{(d)}{=} I_{t-1}\Big(\tilde\theta_t^*; Y_{\alpha(\tilde\theta_t)}|\tilde\theta_t\Big) \\
&= \sum_{i=1}^m \mathbb{P}\Big(\tilde\theta_t = \theta^i\Big) \cdot I_{t-1}\Big(\tilde\theta_t^*; Y_{\alpha(\tilde\theta_t)}|\tilde\theta_t = \theta^i\Big) \\
&= \sum_{i=1}^m \mathbb{P}\Big(\tilde\theta_t^* = \theta^i\Big) \cdot I_{t-1}\Big(\tilde\theta_t^*; Y_{\alpha(\theta^i)}\Big) \\
&= \sum_{i=1}^m \sum_{j=1}^m \mathbb{P}\Big(\tilde\theta_t^* = \theta^i\Big)\mathbb{P}\Big(\tilde\theta_t^* = \theta^j\Big) \cdot D_{\mathrm{KL}}\Big(P_{t-1}\big(Y_{\alpha(\theta^i)}|\tilde\theta_t^* = \theta^j\big) \big\| P_{t-1}\big(Y_{\alpha(\theta^i)}\big)\Big) \\
&\stackrel{(e)}{\geq} 2\sum_{i=1}^m \sum_{j=1}^m \mathbb{P}\Big(\tilde\theta_t^* = \theta^i\Big)\mathbb{P}\Big(\tilde\theta_t^* = \theta^j\Big) \cdot \Big\{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})|\tilde\theta_t^* = \theta^j\big] - \mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big]\Big\}^2,
\end{aligned}$$

where (d) comes from the fact that $\tilde\theta_t^*$ and $\tilde\theta_t$ are independent, conditioned on $\mathcal{H}_{t-1}$, and (e) follows from Pinsker's inequality and our assumption that $\sup_{y\in\mathcal{Y}} R(y) - \inf_{y\in\mathcal{Y}} R(y) \leq 1$.

On the other hand, there is also

$$\begin{aligned}
\mathbb{E}_{t-1}\big[R(Y_{\tilde\theta_t^*}) - R(Y_{\tilde\theta_t})\big] &= \sum_{i=1}^m \mathbb{P}_{t-1}\Big(\tilde\theta_t^* = \theta^i\Big)\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta_i)})|\tilde\theta_t^* = \theta^i\big] - \\
&\quad \sum_{i=1}^m \mathbb{P}_{t-1}\Big(\tilde\theta_t = \theta^i\Big)\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big] \\
&= \sum_{i=1}^m \mathbb{P}_{t-1}\Big(\tilde\theta_t^* = \theta^i\Big)\Big\{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})|\tilde\theta_t^* = \theta^i\big] - \mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big]\Big\}.
\end{aligned}$$

All equalities and inequalities hold almost surely. Thus the proof is complete. □

Lemma 3. *For each $t = 1, 2, \cdots$, there is*

$$\tilde\Gamma_t(\tilde\theta_t^*; \tilde\theta_t) \leq \frac{d}{2}, \quad \text{a.s.}$$

Proof. Fix $t \in \{1, \cdots, T\}$, and let

$$q_i = \mathbb{P}_{t-1}\big(\tilde\theta_t^* = \theta^i\big),\ s_i = \mathbb{E}_{t-1}\big[\theta^*|\tilde\theta_t^* = \theta^i\big], \quad i = 1, \cdots, m,$$

and $s = \mathbb{E}_{t-1}\big[\theta^*\big]$. The linearity of expectation gives us

$$\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})|\tilde\theta_t^* = \theta^j\big] = \alpha(\theta_i)^\top s_j, \quad \mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta^i)})\big] = \alpha(\theta^i)^\top s, \quad \forall i, j \in \{1, \cdots, m\}.$$

From Lemma 2, we have

$$\begin{aligned}
\tilde\Gamma_t(\tilde\theta_t^*, \tilde\theta_t) &= \frac{\mathbb{E}_{t-1}\big[R(Y_{\alpha(\tilde\theta_t^*)}) - R(Y_{\alpha(\tilde\theta_t)})\big]^2}{I_{t-1}\Big(\tilde\theta_t^*; (\tilde\theta_t, Y_{\alpha(\tilde\theta_t)})\Big)} \\
&\leq \frac{\Big(\sum_{i=1}^m q_i(\alpha(\theta^i)^\top s_i - \alpha(\theta^i)^\top s)\Big)^2}{2\sum_{i=1}^m \sum_{j=1}^m q_i q_j (\alpha(\theta^i)^\top s_j - \alpha(\theta^i)^\top s)^2} \\
&= \frac{\Big(\sum_{i=1}^m q_i \alpha(\theta_i)^\top (s_i - s)\Big)^2}{2\sum_{i=1}^m \sum_{j=1}^m q_i q_j \big[\alpha(\theta_i)^\top (s_j - s)\big]^2} \quad \text{a.s.}
\end{aligned}$$

Let $u_i = \sqrt{q_i}\alpha(\theta^i)$ and $v_i = \sqrt{q_i}(s_i - s)$, then $u_i, v_i \in \mathbb{R}^d$ for $i = 1, \cdots, m$. Consider the matrix

$$M = (u_i^\top v_j)_{i,j=1}^m = \begin{pmatrix} u_1^\top \\ u_2^\top \\ \vdots \\ u_m^\top \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix}.$$

Notice that M is the product of an $m \times d$ matrix and a $d \times m$ matrix, hence $\mathrm{rank}(m) \le d$. Therefore we have

$$\tilde{\Gamma}_t(\tilde{\theta}_t^*, \tilde{\theta}_t) \le \frac{\mathrm{Trace}(M)^2}{2\|M\|_{\mathrm{F}}^2} \le \frac{\mathrm{rank}(M)}{2} \le \frac{d}{2}, \quad \text{a.s.}$$

$\square$

Notice that

$$\begin{aligned} \mathbb{E}\big[R(Y_{\alpha(\theta_1)}) - R(Y_{\alpha(\theta_2)})\big]^2 &\le \mathbb{E}\Big[\mathbb{E}_{t-1}\big[R(Y_{\alpha(\theta_1)}) - R(Y_{\alpha(\theta_2)})\big]^2\Big] \\ &= \mathbb{E}\Big[\tilde{\Gamma}_t(\theta_1;\theta_2)\cdot I_{t-1}\big(\theta_1;(\theta_2, Y_{\alpha(\theta_2)})\big)\Big] \\ &\overset{(f)}{\le} \frac{d}{2}\cdot \mathbb{E}\Big[I_{t-1}\big(\theta_1;(\theta_2, Y_{\alpha(\theta_2)})\big)\Big] \\ &= \frac{d}{2}\cdot I\big(\theta_1;(\theta_2, Y_{\alpha(\theta_2)})\big|\tilde{\mathcal{H}}_{t-1}\big), \end{aligned} \tag{A11}$$

where (f) comes from Lemma 3. Hence the proof is complete.

C Proof of Proposition 4

Let $\{\mathcal{A}_k\}_{k=1}^K$ be an 2ϵ-covering of $\mathcal{A}$ with respect to the Euclidean norm, i.e.

$$\|a_1 - a_2\|_2 \le 2\epsilon, \quad \forall a_1, a_2 \in \mathcal{A}_k,\ k = 1, \cdots, K.$$

Define

$$\Theta_k = \alpha^{-1}(\mathcal{A}_k) = \{\theta \in \Theta : \alpha(\theta) \in \mathcal{A}_k\}.$$

Apparently $\{\Theta_k\}_{k=1}^K$ is a partition of Θ. Moreover, for any $k \in \{1, \cdots, K\}$ and $\theta_1, \theta_2 \in \Theta_k$ there is

$$\begin{aligned} \big|\mu(\alpha(\theta_1), \theta_2) - \mu(\alpha(\theta_2), \theta_2)\big| &= \frac{1}{2}\big|\alpha(\theta_1)^\top\theta_2 - \alpha(\theta_2)^\top\theta_2\big| \\ &\le \frac{1}{2}\|\alpha(\theta_1) - \alpha(\theta_2)\|_2 \cdot \|\theta_2\|_2 \le \epsilon, \end{aligned} \tag{A12}$$

where the last inequality follows from that $\|a_1 - a_2\|_2 \le 2\epsilon$ and that $\Theta \subseteq \overline{\mathbf{B}_d(0,1)}$.

Let $N(S, \epsilon, \|\cdot\|)$ be the ϵ-covering number of set S with respect to the $\|\cdot\|$-norm. We only have to bound $N(\mathcal{A}, 2\epsilon, \|\cdot\|_2)$. From a standard result,

$$N(\mathcal{A}, 2\epsilon, \|\cdot\|_2) \le N\Big(\overline{\mathbf{B}_d(0,1)}, 2\epsilon, \|\cdot\|_2\Big) \le \left(\frac{1}{\epsilon} + 1\right)^d.$$

Therefore

$$K \le \left(\frac{1}{\epsilon} + 1\right)^d.$$

D Proof of Theorem 2

From Theorem 1, Propositions 3 and 4, we have that for all $\epsilon > 0$,

$$\mathrm{BayesRegret}(T;\pi^{\mathrm{TS}}) \le \sqrt{\frac{d}{2}\cdot d\log\left(\frac{1}{\epsilon}+1\right)\cdot T} + \epsilon\cdot T.$$

Taking $\epsilon = d/\sqrt{2T}$, we arrive at

$$\begin{aligned} \mathrm{BayesRegret}(T;\pi^{\mathrm{TS}}) &\le d\sqrt{\frac{T}{2}}\left(\sqrt{\log\left(1+\frac{\sqrt{2T}}{d}\right)}+1\right) \\ &\le d\sqrt{T}\cdot\sqrt{\log\left(1+\frac{\sqrt{2T}}{d}\right)+1} \\ &\le d\sqrt{T\log\left(3+\frac{3\sqrt{2T}}{d}\right)}. \end{aligned}$$

E Proof of Propositions 5, 6 and Theorem 3

Let W be a random variable with the same distribution as the noise W_a for all $a \in \mathcal{A}$. Define function f as

$$f(x) = \mathbb{E}\big[\phi^{-1}(x - W)\big].$$

For each $a \in \mathcal{A}$, let $S_a = f(R_a) = \mathbb{E}\big[\phi^{-1}(R_a - W)\big|R_a\big]$. Then we have

$$\begin{aligned}
\mathbb{E}\big[S_a\big|\theta^* = \theta\big] &= \mathbb{E}\Big[\mathbb{E}\big[\phi^{-1}(R_a - W)\big|R_a\big]\Big|\theta^* = \theta\Big] \\
&\overset{(g)}{=} \mathbb{E}\Big[\mathbb{E}\big[\phi^{-1}(R_a - W_a)\big|R_a\big]\Big|\theta^* = \theta\Big] \\
&\overset{(h)}{=} \mathbb{E}\Big[\mathbb{E}\big[a^\top\theta\big|R_a\big]\Big|\theta^* = \theta\Big] \\
&= a^\top\theta, \qquad \text{(A13)}
\end{aligned}$$

where (g) follows from the fact that W and W_a have the same distribution, and (h) results from that conditioned on $p^* = p$,

$$R_a = \phi(a^\top\theta) + W_a.$$

From Lemma 3, we have that

$$\frac{\mathbb{E}_{t-1}\big[S_{\alpha(\tilde\theta_t^*)} - S_{\alpha(\tilde\theta_t)}\big]^2}{I_{t-1}\big(\tilde\theta_t^*; (\tilde\theta_t, S_{\alpha(\tilde\theta_t)})\big)} \le 2d.$$

Notice that the constant is different from that in Lemma 3 since we have $S_a \in [-1, 1]$ for all $a \in \mathcal{A}$, whereas in Lemma 3 there is $R_a \in [-1/2, 1/2]$. From data-processing inequality, since $S_{\alpha(\tilde\theta_t)} = f\big(R_{\alpha(\tilde\theta_t)}\big)$, there should be

$$I\big(\tilde\theta_t^*; (\tilde\theta_t, S_{\alpha(\tilde\theta_t)})\big) \le I\big(\tilde\theta_t^*; (\tilde\theta_t, R_{\alpha(\tilde\theta_t)})\big).$$

Also there is

$$\begin{aligned}
\mathbb{E}\big[S_{\alpha(\tilde\theta_t^*)} - S_{\alpha(\tilde\theta_t)}\big]^2 &= \mathbb{E}\big[f\big(R_{\alpha(\tilde\theta_t^*)}\big) - f\big(R_{\alpha(\tilde\theta_t)}\big)\big]^2 \\
&\ge \big[\inf_x f'(x)\big]^2 \cdot \mathbb{E}\big[R_{\alpha(\tilde\theta_t^*)} - R_{\alpha(\tilde\theta_t)}\big]^2 \\
&\overset{(i)}{\ge} C(\phi)^{-2} \cdot \mathbb{E}\big[R_{\alpha(\tilde\theta_t^*)} - R_{\alpha(\tilde\theta_t)}\big]^2,
\end{aligned}$$

where (i) is the consequence of

$$\inf_x f'(x) = \inf_x \mathbb{E}\big[(\phi^{-1})'(x - W)\big] \ge \inf_x (\phi^{-1})'(x) = \left[\sup_{y\in[-1,1]} \phi'(y)\right]^{-1}.$$

Therefore there is

$$\tilde\Gamma_t(\tilde\theta_t^*; \tilde\theta_t) \le 2C(\phi)^2 d$$

where $\tilde\Gamma$ is defined in (A10). This proves Proposition 5.

On the other hand, let $\{\mathcal{A}_k\}_{k=1}^K$ be an $\epsilon/C(\phi)$-covering of $\mathcal{A}$ with respect to the Euclidean norm, i.e.

$$\|a_1 - a_2\|_2 \le \epsilon/C(\phi), \quad \forall a_1, a_2 \in \mathcal{A}_k,\ k = 1, \cdots, K.$$

Define

$$\Theta_k = \alpha^{-1}(\mathcal{A}_k) = \{\theta \in \Theta : \alpha(\theta) \in \mathcal{A}_k\}.$$

Apparently $\{\Theta_k\}_{k=1}^K$ is a partition of Θ. Moreover, for any $k \in \{1, \cdots, K\}$ and $\theta_1, \theta_2 \in \Theta_k$ there is

$$\begin{aligned}
\big|\mu(\alpha(\theta_1), \theta_2) - \mu(\alpha(\theta_2), \theta_2)\big| &= \big|\phi(\alpha(\theta_1)^\top\theta_2) - \phi(\alpha(\theta_2)^\top\theta_2)\big| \\
&= C(\phi) \cdot \big|\alpha(\theta_1)^\top\theta_2 - \alpha(\theta_2)^\top\theta_2\big| \\
&\le C(\phi) \cdot \|\alpha(\theta_1) - \alpha(\theta_2)\|_2 \cdot \|\theta_2\|_2 \le \epsilon, \qquad \text{(A14)}
\end{aligned}$$

where the last inequality follows from that $\|a_1 - a_2\|_2 \le \epsilon/C(\phi)$ and that $\Theta \subseteq \overline{\mathbf{B}_d(0, 1)}$.

Similar as in the proof of Proposition 4,

$$N(\mathcal{A}, \epsilon/C(\phi), \|\cdot\|_2) \le N\Big(\overline{\mathbf{B}_d(0,1)}, \epsilon/C(\phi), \|\cdot\|_2\Big) \le \left(\frac{2C(\phi)}{\epsilon}+1\right)^d.$$

Therefore

$$K \le \left(\frac{2C(\phi)}{\epsilon}+1\right)^d.$$

Therefore by choosing $\epsilon = \sqrt{2}C(\phi)d/\sqrt{T}$ in Theorem 1, we arrive at

$$\begin{aligned}
\mathrm{BayesRegret}(T;\pi^{\mathrm{TS}}) &\le \sqrt{2C(\phi)^2 d\cdot d\log\left(1+\frac{\sqrt{2T}}{d}\right)\cdot T} + \sqrt{2}C(\phi)d\sqrt{T} \\
&\le \sqrt{2}C(\phi)d\sqrt{T}\left(\sqrt{\log\left(1+\frac{\sqrt{2T}}{d}\right)}+1\right) \\
&\le 2C(\phi)d\sqrt{T}\cdot\sqrt{\log\left(1+\frac{\sqrt{2T}}{d}\right)+1} \\
&\le 2C(\phi)d\sqrt{T\log\left(3+\frac{3\sqrt{2T}}{d}\right)}.
\end{aligned}$$

F Proof of Theorem 4

For simplicity, we omit the superscript L in ϕ^{L} throughout this proof. We first show that, for any $\epsilon \in (0, \phi(\delta) - 1/2)$ there exists a partition $\{\Theta_k\}_{k=1}^K$ such that (3) holds and

$$K \le \frac{1}{\epsilon}\left(1+\frac{2}{\delta-\phi^{-1}(\phi(\delta)-\epsilon)}\right)^d. \tag{A15}$$

Let real-number sequence $s_0, s_1, \cdots, s_L$ be defined by

$$\begin{aligned}
s_0 &= \phi^{-1}(\phi(\delta)-\epsilon), \\
s_1 &= \delta, \\
s_2 &= \phi^{-1}(\phi(\delta)+\epsilon), \\
s_3 &= \phi^{-1}(\phi(\delta)+2\epsilon), \\
&\cdots \\
s_{L-1} &= \phi^{-1}(\phi(\delta)+(L-2)\epsilon), \\
s_L &= 1,
\end{aligned}$$

where we choose L such that $\phi(\delta)+(L-2)\epsilon < \phi(1) \le \phi(\delta)+(L-1)\epsilon$. In addition, let $s'_j = -s_j$ for $j = 0, \dots, L$. Notice that since $0 < \epsilon < \phi(\delta) - 1/2$, we have $s_0 > 0$. For $\ell = 1, \cdots, L-1$, let

$$\mathcal{Q}_\ell = \big\{\theta \in \Theta : s_\ell < \alpha(\theta)^\top\theta \le s_{\ell+1}\big\},$$

and let

$$\mathcal{Q}_0 = \big\{\theta \in \Theta : s_0 \le \alpha(\theta)^\top\theta \le s_1\big\}.$$

Similarly for $\ell = 1, \cdots, L-1$, we can define

$$\mathcal{Q}'_\ell = \big\{\theta \in \Theta : s'_{\ell+1} < \alpha(\theta)^\top\theta \le s'_\ell\big\},$$

and

$$\mathcal{Q}'_0 = \big\{\theta \in \Theta : s'_1 \le \alpha(\theta)^\top\theta \le s'_0\big\}.$$

From our assumption there is $\delta \le |\alpha(\theta)^\top \theta| \le 1$ for all $\theta \in \Theta$, hence

$$\left(\bigcup_{\ell=0}^{L-1} \mathcal{Q}_\ell\right) \cup \left(\bigcup_{\ell=0}^{L-1} \mathcal{Q}'_\ell\right) = \Theta.$$

For each $\ell = 1, \ldots, L$, let $\{\mathcal{A}_{\ell j}\}_{j=1}^{J_\ell}$ be an $(s_\ell - s_{\ell-1})$-covering of $\mathcal{A}$ with respect to the Euclidean norm, i.e. for each $j = 1, \cdots, J_\ell$,

$$\|a_1 - a_2\|_2 \le s_\ell - s_{\ell-1}, \quad \forall a_1, a_2 \in \mathcal{A}_{\ell j}.$$

And let $\{\mathcal{A}'_{\ell j}\}_{j=1}^{J'_\ell}$ be an $(s'_{\ell-1} - s'_\ell)$-covering of $\mathcal{A}$ with respect to the Euclidean norm. Correspondingly, let $\{\Theta_{\ell j}\}_{j=1}^{J_\ell}$ be defined by

$$\Theta_{\ell j} = \big\{\theta \in \mathcal{Q}_\ell : \alpha(\theta) \in \mathcal{A}_{\ell j}\big\}.$$

Then $\{\Theta_{\ell j}\}_{j=1}^{J_\ell}$ is a partition of $\mathcal{Q}_\ell$, and for each $j = 1, \cdots, J_\ell$, let $\theta, \theta' \in \Theta_{\ell j}$, there is

$$\begin{aligned}
\mu(\alpha(\theta), \theta) - \mu(\alpha(\theta'), \theta) &= \phi\big(\alpha(\theta)^\top \theta\big) - \phi\big(\alpha(\theta')^\top \theta\big) \\
&\overset{(j)}{\le} \phi\big(\alpha(\theta)^\top \theta\big) - \phi\big(\alpha(\theta)^\top \theta - (s_\ell - s_{\ell-1})\big) \\
&\overset{(k)}{\le} \phi\big(s_\ell\big) - \phi\big(s_\ell - (s_\ell - s_{\ell-1})\big) = \epsilon,
\end{aligned}$$

where (j) comes from that

$$\alpha(\theta)^\top \theta - \alpha(\theta')^\top \theta \le \|\alpha(\theta) - \alpha(\theta')\|_2 \|\theta\|_2 \le s_\ell - s_{\ell-1},$$

and (k) follows from the fact that $\phi\big(x\big) - \phi\big(x - (s_\ell - s_{\ell-1})\big)$ is decreasing in x when $x > s_\ell - s_{\ell-1}$. Let $\{\Theta'_{\ell j}\}_{\ell,j}$ be the counterpart of $\{\Theta_{\ell j}\}_{\ell,j}$ defined with respect to $\{\mathcal{A}'_{\ell j}\}_{\ell,j}$, then $\{\Theta_{\ell j}\}_{\ell,j} \cup \{\Theta'_{\ell j}\}_{\ell,j}$ is a valid partition of Θ. Notice that

$$s_1 - s_0 < s_2 - s_1 < \cdots < s_L - s_{L-1},$$

we thence have

$$\begin{aligned}
K &\le \sum_{\ell=1}^{L} J_\ell + \sum_{\ell=1}^{L} J'_\ell \\
&\le 2\sum_{\ell=1}^{L} N(\mathcal{A}, s_\ell - s_{\ell-1}, \|\cdot\|_2) \\
&\le 2L \cdot N(\mathcal{A}, s_1 - s_0, \|\cdot\|_2) \\
&\le \frac{1}{\epsilon}\left(1 + \frac{2}{\delta - \phi^{-1}(\phi(\delta) - \epsilon)}\right)^d. \qquad \text{(A16)}
\end{aligned}$$

Hence we have proved (A15). Let $\Phi(x) = \delta - \phi^{-1}(\phi(\delta) - x)$, then there is $\Phi(0) = 0$ and $\Phi'(0) = \frac{1}{\phi(\delta)(1-\phi(\delta))}$. Also notice that

$$\begin{aligned}
\Phi''(0) &= -(\phi^{-1})''(\phi(\delta) - x)\Big|_{x=0} \\
&= \frac{2\phi(\delta) - 1}{(\phi(\delta) - \phi(\delta)^2)^2} > 0, \qquad \text{(A17)}
\end{aligned}$$

where we used the fact that $\phi(\delta) > 1/2$. Hence for small enough ϵ, there is $\Phi(\epsilon) \ge \Phi'(0) \cdot \epsilon$. Notice that from Theorem 1 and Conjecture 1, for all T and $\epsilon > 0$ there is

$$\begin{aligned}
\mathrm{BayesRegret}(T; \pi^{\mathrm{TS}}) &\le \sqrt{\frac{d}{2} \cdot \log K \cdot T} + \epsilon \cdot T \\
&\le \sqrt{\frac{d}{2} \cdot \left(-\log(\epsilon) + d \log\left(1 + \frac{2}{\Phi(\epsilon)}\right)\right) \cdot T} + \epsilon \cdot T \qquad \text{(A18)}
\end{aligned}$$

Let $\epsilon = d/\sqrt{2T}$, for large enough T we have

$$\begin{aligned}
\mathrm{BayesRegret}(T;\pi^{\mathrm{TS}}) &\leq \sqrt{\frac{d}{2}\cdot\left(\log(\frac{\sqrt{2T}}{d}) + d\log\left(1+\frac{2\sqrt{2T}}{\Phi'(0)d}\right)\right)\cdot T} + d\sqrt{\frac{T}{2}} \\
&\leq \sqrt{\frac{d(d+1)}{2}\cdot T}\cdot\left(\sqrt{\log\left(1+\frac{2\sqrt{2T}}{\Phi'(0)d}\right)}+1\right) \\
&\leq \sqrt{d(d+1)T}\cdot\sqrt{\log\left(1+\frac{2\sqrt{2T}}{\Phi'(0)d}\right)+1} \\
&\leq 2d\sqrt{T}\cdot\sqrt{\log\left(3+\frac{6\sqrt{2T}}{d}\cdot\frac{\beta e^{\beta\delta}}{(1+e^{\beta\delta})^2}\right)}. \qquad \text{(A19)}
\end{aligned}$$