Supplementary Material

Mixed Strategies for Robust Optimization of Unknown Objectives

Pier Giuseppe Sessa, Ilija Bogunovic, Maryam Kamgarpour, Andreas Krause (AISTATS 2020)

A Proof of Theorem 2

Proof. In this proof, we condition on the event in Lemma 1 holding true, meaning that ucb_t and lcb_t provide valid confidence bounds as per (13). As stated in the lemma, this holds with probability at least $1-\delta$.

Our main goal in this proof is to upper bound the difference:

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] - \min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} f(\boldsymbol{x}_t,\boldsymbol{\theta}). \tag{19}$$

To do so, we provide upper and lower bounds of the first and second terms, respectively, and then we upper bound their difference.

First, we show that the following holds:

$$\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} f(\boldsymbol{x}_t,\boldsymbol{\theta}) \geq \left(\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - 4\beta_T\sqrt{\frac{\lambda\gamma_T}{T}}, \tag{20}$$

where $\boldsymbol{x}_t$ is the point queried at time t.

To prove Eq. (20) we use the lower confidence bound and (14):

$$\begin{aligned}
\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} f(\boldsymbol{x}_t,\boldsymbol{\theta}) &\geq \min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \mathrm{lcb}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) && (21)\\
&= \min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \big(\mathrm{ucb}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) - 2\beta_t\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\big) && (22)\\
&\geq \left(\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - \max_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} 2\beta_t\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) && (23)\\
&\geq \left(\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - \frac{2\beta_T}{T}\sum_{t=1}^{T} \max_{\boldsymbol{\theta}\in\Theta} \sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) && (24)\\
&= \left(\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - \frac{2\beta_T}{T}\sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_t) && (25)\\
&\geq \left(\min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}, && (26)
\end{aligned}$$

where (22) follows from the definition of the confidence bounds in (5) and (6), (24) is due to monotonicty of β_t, and (25) is by rule (10) used in Algorithm 1 to select $\boldsymbol{\theta}_t$. Finally, (26) is obtained via the standard result from (Srinivas et al., 2010; Chowdhury and Gopalan, 2017)

$$\sum_{t=1}^{T} \sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_t) \leq \sqrt{4T\lambda\gamma_T}, \tag{27}$$

when $\lambda \geq 1$.

Next, we show that the first term can be upper bounded as follows:

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] \leq \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}_{\boldsymbol{\theta}\sim\boldsymbol{w}_t}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})].$$

To prove this, we start by upper bounding the minimum value of the inner objective:

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})}\min_{\boldsymbol{\theta}\in\Theta}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] \leq \max_{\mathcal{P}\in\Delta(\mathcal{X})}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta}_i)] \tag{28}$$

$$\leq \max_{\mathcal{P}\in\Delta(\mathcal{X})}\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}\left[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta}_i)\right] \tag{29}$$

$$= \max_{\mathcal{P}\in\Delta(\mathcal{X})}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}\left[\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta}_i)\right] \tag{30}$$

$$\leq \frac{1}{T}\sum_{t=1}^{T}\max_{\mathcal{P}\in\Delta(\mathcal{X})}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}\left[\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta}_i)\right] \tag{31}$$

$$= \frac{1}{T}\sum_{t=1}^{T}\max_{\boldsymbol{x}\in\mathcal{X}}\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta}_i) \tag{32}$$

$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_i)\,. \tag{33}$$

We obtain Eq. (28) as the following trivially holds

$$\min_{\boldsymbol{\theta}\in\Theta}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] \leq \sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta}_i)]$$

for each t and $\boldsymbol{w}_t \in \{\boldsymbol{w}\in[0,1]^m : \sum_{i=1}^m \boldsymbol{w}[i]=1\}$, and hence it also holds for the average value

$$\min_{\boldsymbol{\theta}\in\Theta}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] \leq \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m}\boldsymbol{w}_t[i]\cdot\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta}_i)].$$

Eq. (29) follows from (14), (30) follows by the linearity of expectation, and (32) holds since Dirac delta $\boldsymbol{\delta}_{\boldsymbol{x}}$, $\forall\boldsymbol{x}\in\mathcal{X}$, is in $\Delta(\mathcal{X})$. Finally, (33) follows by rule (9) used in Algorithm 1 to select $\boldsymbol{x}_t$.

Next, we bound the difference in (19) by combining the bounds obtained in (26) and (33):

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})}\min_{\boldsymbol{\theta}\in\Theta}\mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] - \min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}f(\boldsymbol{x}_t,\boldsymbol{\theta})$$
$$\leq \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\boldsymbol{\theta}\sim\boldsymbol{w}_t}\left[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right] - \left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) + 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}$$
$$\leq \sqrt{\frac{\log(m)}{2T}} + 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}, \tag{34}$$

where (34) follows by the guarantees of the no-regret online multiplicative weight updates algorithm played by the adversary, that is,

$$\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_{\boldsymbol{\theta}\sim\boldsymbol{w}_t}\left[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right] - \left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) \leq \sqrt{\frac{\log(m)}{2T}}, \tag{35}$$

with the learning rate set to $\eta_T = \sqrt{\frac{8\log(m)}{T}}$. For more details on this result see (Cesa-Bianchi and Lugosi, 2006, Section 4.2) where the same online algorithm is considered. Specifically, the result above follows from (Cesa-Bianchi and Lugosi, 2006, Theorem 2.2) by noting that $\sum_{t=1}^{T}\mathbb{E}_{\boldsymbol{\theta}\sim\boldsymbol{w}_t}\left[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right] = \sum_{t=1}^{T}\boldsymbol{w}_t^T\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\cdot)$, $\min_{\boldsymbol{\theta}\in\Theta}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) = \min_{\boldsymbol{w}\in\Delta(\Theta)}\sum_{t=1}^{T}\boldsymbol{w}^T\cdot\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\cdot)$ and $\overline{\mathrm{ucb}}_{t-1}(\cdot,\cdot)\in[0,1]$ for every t. In our case, the objective function changes with t but remains bounded, which allows the result to hold despite the changes (see time-varying games result extension (Cesa-Bianchi and Lugosi, 2006, Remark 7.3)).

By rearranging (34) and by letting $\mathcal{U}^{(T)}$ be the uniform distribution over the queried points $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_T\}$ during the run of Algorithm 1, we obtain:

$$\min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{U}^{(T)}}[f(\boldsymbol{x},\boldsymbol{\theta})] \geq \max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[f(\boldsymbol{x},\boldsymbol{\theta})] - \sqrt{\frac{\log(m)}{2T}} - 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}.$$

Finally, we require $\epsilon \geq \sqrt{\frac{\log(m)}{2T}} + 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}$, which we obtain when

$$T \geq \frac{1}{\epsilon^2}\left(\frac{\log(m)}{2} + \beta_T\sqrt{32\lambda\gamma_T\log(m)} + 16\beta_T^2\lambda\gamma_T\right).$$

□

B Proof of Corollary 3

Proof. The proof closely follows the one of Theorem 2. The main changes are due to the modified best-response rule from (16).

For a given distribution $\mathcal{Q}\in\Delta(\Theta)$ and trade-off parameter $\chi\in(0,1]$, we can define the new function

$$g(\boldsymbol{x},\boldsymbol{\theta}) := \chi\cdot f(\boldsymbol{x},\boldsymbol{\theta}) + (1-\chi)\cdot\mathbb{E}_{\boldsymbol{\theta}\sim\mathcal{Q}}[f(\boldsymbol{x},\boldsymbol{\theta})] \tag{36}$$

Same as before, our goal is to upper bound the difference:

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta})] - \min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{x}_t,\boldsymbol{\theta}), \tag{37}$$

where $\boldsymbol{x}_t$ is the point selected at time t by GP-MRO using the modified best-response rule as in (16).

Next, we condition on the event in Lemma 1 holding true, and we provide upper and lower bounds of the first and second term, respectively.

First, we show that the second term of (37) can be lower bounded as:

$$\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{x}_t,\boldsymbol{\theta}) \geq \chi\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right) - 4\beta_T\sqrt{\frac{\lambda\gamma_T}{T}}. \tag{38}$$

To prove Eq. (38) we make use of (36) and similar arguments as the ones used in the proof of Theorem 2:

$$\begin{aligned}
&\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{x}_t,\boldsymbol{\theta}) = \chi\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T} f(\boldsymbol{x}_t,\boldsymbol{\theta})\right) + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[f(\boldsymbol{x}_t,\boldsymbol{\theta})]\right)\\
&\geq \chi\left[\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - \frac{2\beta_T}{T}\sum_{t=1}^{T}\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_t)\right] + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[f(\boldsymbol{x}_t,\boldsymbol{\theta})]\right)\\
&\geq \chi\left[\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) - \frac{2\beta_T}{T}\sum_{t=1}^{T}\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_t)\right] + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}) - 2\beta_t\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right)\\
&\geq \chi\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right) - \frac{2\beta_T}{T}\sum_{t=1}^{T}\sigma_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta}_t)\\
&\geq \chi\left(\min_{\boldsymbol{\theta}\in\Theta}\frac{1}{T}\sum_{t=1}^{T}\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\right) + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right) - 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}.
\end{aligned}$$

Next, we show that the first term of (37) can be upper bounded as:

$$\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta})] \leq \chi\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\boldsymbol{w}_t}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right) + (1-\chi)\left(\frac{1}{T}\sum_{t=1}^{T}\mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\right). \tag{39}$$

To prove this we use similar arguments as in the proof of Theorem 2:

$$
\begin{aligned}
\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta})] &\le \max_{\mathcal{P}\in\Delta(\mathcal{X})} \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta}_i)] \\
&\le \frac{1}{T}\sum_{t=1}^{T} \max_{\mathcal{P}\in\Delta(\mathcal{X})} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}\Big[\sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot g(\boldsymbol{x},\boldsymbol{\theta}_i)\Big] \\
&= \frac{1}{T}\sum_{t=1}^{T} \max_{\boldsymbol{x}\in\mathcal{X}} \sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot g(\boldsymbol{x},\boldsymbol{\theta}_i) \\
&= \frac{1}{T}\sum_{t=1}^{T} \max_{\boldsymbol{x}\in\mathcal{X}} \Big[\chi\cdot\sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot f(\boldsymbol{x},\boldsymbol{\theta}_i) + (1-\chi)\cdot \mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[f(\boldsymbol{x},\boldsymbol{\theta})]\Big] \\
&\le \frac{1}{T}\sum_{t=1}^{T} \max_{\boldsymbol{x}\in\mathcal{X}} \Big[\chi\cdot\sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta}_i) + (1-\chi)\cdot \mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x},\boldsymbol{\theta})]\Big] \\
&= \chi\Big(\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{m} \boldsymbol{w}_t[i]\cdot \overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})\Big) + (1-\chi)\Big(\frac{1}{T}\sum_{t=1}^{T} \mathop{\mathbb{E}}_{\boldsymbol{\theta}\sim\mathcal{Q}}[\overline{\mathrm{ucb}}_{t-1}(\boldsymbol{x}_t,\boldsymbol{\theta})]\Big), \qquad (40)
\end{aligned}
$$

where (40) is obtained by the rule in (16) used to select $\boldsymbol{x}_t$.

Next, we bound the difference in (37) by combining the bounds (38) and (39) and applying (35) to obtain:

$$
\max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta})] - \min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{x}_t,\boldsymbol{\theta}) \le \chi\sqrt{\frac{\log(m)}{2T}} + 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}, \qquad (41)
$$

By letting $\mathcal{U}^{(T)}$ be the uniform distribution over the queried points $\{\boldsymbol{x}_1,\dots,\boldsymbol{x}_T\}$ and by using the definitions of $W(\cdot)$ and $\mathcal{P}^*$ together with the bound (41), we obtain:

$$
\begin{aligned}
W(\mathcal{U}^{(T)}) = \min_{\boldsymbol{\theta}\in\Theta} \frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{x}_t,\boldsymbol{\theta}) &\ge \max_{\mathcal{P}\in\Delta(\mathcal{X})} \min_{\boldsymbol{\theta}\in\Theta} \mathbb{E}_{\boldsymbol{x}\sim\mathcal{P}}[g(\boldsymbol{x},\boldsymbol{\theta})] - \chi\sqrt{\frac{\log(m)}{2T}} - 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}} \\
&= W(\mathcal{P}^*) - \chi\sqrt{\frac{\log(m)}{2T}} - 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}} \qquad (42)
\end{aligned}
$$

Finally, we require $\epsilon \ge \chi\sqrt{\frac{\log(m)}{2T}} + 4\beta_T\sqrt{\frac{\gamma_T\lambda}{T}}$, which we obtain when

$$
T \ge \frac{1}{\epsilon^2}\Big(\frac{\chi^2\log(m)}{2} + \chi\,\beta_T\sqrt{32\lambda\gamma_T\log(m)} + 16\beta_T^2\lambda\gamma_T\Big).
$$

□