E-companion: Additional Proofs for Section 4

Appendix EC.1: Proof of Theorem 1

Proof of Theorem 1. Suppose d_1 is the underlying demand function. The regret under demand d_1 can be decomposed as

$$\text{Regret}_1^{\mathsf{mPC}}(T) = \mathbb{E}_1\left[\sum_{t=1}^{T}(r_1^* - r_1(P_t))\right] = \sum_{\ell=0}^{m}\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right].$$

We first consider the case where $\log^{(m)} T > 0$. By definition, $\tau_1 = \left\lceil M_\Phi(P_0^*)\log^{(m)} T\right\rceil$, so the regret during Phase 0 is equal to

$$\mathbb{E}_1\left[\sum_{t=1}^{\tau_1}(r_1^* - r_1(P_t))\right] = \left\lceil M_\Phi(P_0^*)\log^{(m)} T\right\rceil(r_1^* - r_1(P_0^*)). \tag{EC.1}$$

Next, we show that for each $1 \le \ell \le m$, the regret during Phase ℓ is bounded by

$$\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right] \le \frac{2M_\Phi^* r_1^*}{\log^{(m-\ell)} T} + \frac{2r_1^*}{(\log^{(m-\ell)} T)^2}, \tag{EC.2}$$

where $M_\Phi^* = \max_{i\in\{1,\cdots,K\}} M_\Phi(p_i^*)$.

For $1 \le \ell \le m$, the regret during Phase ℓ satisfies the following bound:

$$\begin{aligned}
&\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right]\\
=&\mathbb{E}_1\left[(\tau_{\ell+1} - \tau_\ell)\times(r_1^* - r_1(P_\ell^*))\right]\\
\le&\mathbb{E}_1\left[\Big(M_\Phi(P_\ell^*)\log^{(m-\ell)} T + 1\Big)\times(r_1^* - r_1(P_\ell^*))\right]\\
\le&\Big(M_\Phi^*\log^{(m-\ell)} T + 1\Big)\sum_{i=1}^{K}(r_1^* - r_1(p_i^*))\times\mathbb{P}_1[P_\ell^* = p_i^*]\\
\le&\Big(M_\Phi^*\log^{(m-\ell)} T + 1\Big)\, r_1^*\times\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^* = p_i^*].
\end{aligned} \tag{EC.3}$$

In the expression above, the expectation is taken on the price offered in Phase ℓ, P_ℓ^*, which is a random variable depending on the realized demand in phases $0,\cdots,\ell-1$. In Eq (EC.3), we use the fact that for all $\ell = 1,\cdots,m$, the offered price $P_\ell^* \in \{p_1^*,\cdots,p_K^*\}$ (see line 10 of mPC).

To complete the proof of inequality (EC.2), we prove the following inequality:

$$\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^* = p_i^*] \le \frac{2}{(\log^{(m-\ell)} T)^2}. \tag{EC.4}$$

By the definition of mPC, the choice of price P_ℓ^* is determined by the sample mean $\bar{X}_{\ell-1}$ in Phase $\ell-1$, so we have

$$\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^*=p_i^*]=\mathbb{P}_1\left[|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)|\geq|\bar{X}^{\ell-1}-d_i(P_{\ell-1}^*)| \text{ for some } i\neq 1\right].$$

Now, if $|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)|\geq|\bar{X}^{\ell-1}-d_i(P_{\ell-1}^*)|$ for some $i\neq 1$, we have

$$|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)|\geq\frac{1}{2}\left(|\bar{X}^{\ell-1}-d_i(P_{\ell-1}^*)|+|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)|\right)\geq\frac{1}{2}|d_i(P_{\ell-1}^*)-d_1(P_{\ell-1}^*)|,$$

where the last step uses the triangle inequality. This leads to the following bound:

$$\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^*=p_i^*]\leq\mathbb{P}_1\left[\left|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)\right|\geq\frac{1}{2}\min_{i\neq 1}\left|d_1(P_{\ell-1}^*)-d_i(P_{\ell-1}^*)\right|\right]. \tag{EC.5}$$

Given price $P_{\ell-1}^*$, sample mean $\bar{X}_{\ell-1}$ is the average of i.i.d. random variables with mean $d_1(P_{\ell-1})$. Because demand in each period is lighted-tailed with parameters (σ,b), we can apply the Chernoff inequality: conditioning on $P_{\ell-1}^*$, for any $\epsilon>0$, it holds that

$$\mathbb{P}_1\left[|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)|\geq\epsilon\middle|P_{\ell-1}^*\right)\leq 2\exp\left(-(\tau_\ell-\tau_{\ell-1})\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right].$$

Let $\epsilon=\frac{1}{2}\min_{i\neq 1}\left|d_1(P_{\ell-1}^*)-d_i(P_{\ell-1}^*)\right|$. Because $\tau_\ell-\tau_{\ell-1}=\left\lceil M_\Phi(P_{\ell-1}^*)\log^{(m-\ell+1)}T\right\rceil$, we have

$$\begin{aligned}
&\mathbb{P}_1\left[\left|\bar{X}^{\ell-1}-d_1(P_{\ell-1}^*)\right|\geq\frac{1}{2}\min_{i\neq 1}\left|d_1(P_{\ell-1}^*)-d_i(P_{\ell-1}^*)\right|\middle|P_{\ell-1}^*\right]\\
\leq&2\mathbb{E}_1\left[\exp\left(-\left\lceil M_\Phi(P_{\ell-1}^*)\log^{(m-\ell+1)}T\right\rceil\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right)\middle|P_{\ell-1}^*\right]\\
=&2\exp\left(-\left\lceil M_\Phi(P_{\ell-1}^*)\log^{(m-\ell+1)}T\right\rceil\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right)\\
\leq&2\exp\left(-M_\Phi(P_{\ell-1}^*)\log^{(m-\ell+1)}T\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right)\\
\leq&2\exp\left(-2\log^{(m-\ell+1)}T\right) &\text{(EC.6)}\\
=&\frac{2}{\left(\log^{(m-\ell)}T\right)^2},
\end{aligned}$$

where step (EC.6) uses the definition

$$M_\Phi(P_{\ell-1}^*)=2\times\left(\frac{2\sigma^2}{\frac{1}{4}\min_{i\neq j}\left(d_i(P_{\ell-1}^*)-d_j(P_{\ell-1}^*)\right)^2}\vee\frac{2b}{\frac{1}{2}\min_{i\neq j}\left|d_i(P_{\ell-1}^*)-d_j(P_{\ell-1}^*)\right|}\right).$$

By integrating over the realizations of $P_{\ell-1}^*$ in the above bound, we have established inequality (EC.4), which in turn proves (EC.2).

Combining Eqs (EC.1) and (EC.2), we can prove the regret bound on mPC under demand d_1 as follows:

$$\begin{aligned}\text{Regret}_1^{\mathsf{mPC}}(T) &= \mathbb{E}_1\left[\sum_{t=1}^{\tau_1}(r_1^* - r_1(P_t))\right] + \sum_{\ell=1}^{m}\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right]\\ &\le \left(M_\Phi(P_0^*)\log^{(m)}T + 1\right)(r_1^* - r_1(P_0^*)) + \sum_{\ell=1}^{m}\left(\frac{2M_\Phi^* r_1^*}{\log^{(m-\ell)}T} + \frac{2r_1^*}{(\log^{(m-\ell)}T)^2}\right).\end{aligned}$$

Since $\log^{(m)}T > 0$, it is easily verified that $\log^{(m-\ell)}T \ge e^{\ell-1}$ for all $\ell \ge 1$, so

$$\sum_{\ell=1}^{m}\frac{1}{\log^{(m-\ell)}T} \le \sum_{\ell=1}^{\infty}\frac{1}{e^{\ell-1}} \le 2, \quad \sum_{\ell=1}^{m}\frac{1}{(\log^{(m-\ell)}T)^2} \le \sum_{\ell=1}^{\infty}\frac{1}{e^{2\ell-2}} \le \frac{3}{2}.$$

Therefore,

$$\begin{aligned}\text{Regret}_1^{\mathsf{mPC}}(T) &\le \left(M_\Phi(P_0^*)\log^{(m)}T + 1\right)(r_1^* - r_1(P_0^*)) + 4M_\Phi^* r_1^* + 3r_1^*\\ &\le M_\Phi(P_0^*)(r_1^* - r_1(P_0^*))\log^{(m)}T + 4M_\Phi^* r_1^* + 4r_1^*.\end{aligned}$$

The minimax regret of demand set Φ is bounded by

$$\text{Regret}_\Phi^{\mathsf{mPC}}(T) = \max_{i=1,\dots,K}\text{Regret}_i^{\mathsf{mPC}}(T) \le C_\Phi(P_0^*)\log^{(m)}T + 4M_\Phi^* r^* + 4r^*,$$

where $C_\Phi(P_0^*) = \max_{i\in\{1,\cdots,K\}}\{M_\Phi(P_0^*)(r_i^* - r_i(P_0^*))\}$ and $r^* = \max_{i\in\{1,\cdots,K\}} r_i^*$.

If $\log^{(m)}T = 0$, let $m' \le m$ be the largest integer such that $\log^{(m')}T > 0$. Clearly, $\log^{(m')}T \le 1$. In this case, policy mPC applied to T periods uses only m' price changes, so

$$\text{Regret}_\Phi^{\mathsf{mPC}}(T) \le C_\Phi(P_0^*)\log^{(m')}T + 4M_\Phi^* r^* + 4r^* \le C_\Phi(P_0^*) + 4M_\Phi^* r^* + 4r^*.$$

Combining both cases for $\log^{(m)}T > 0$ and $\log^{(m)}T = 0$, we have

$$\text{Regret}_\Phi^{\mathsf{mPC}}(T) \le C_\Phi(P_0^*)\max\{\log^{(m)}T, 1\} + 4M_\Phi^* r^* + 4r^*.$$

$\square$

Appendix EC.2: Proof of Proposition 1

Proof of Proposition 1. Let m be the integer such that $\tau_m < T \le \tau_{m+1}$. Suppose d_1 is the underlying demand function. The regret under demand d_1 can be composed as

$$\text{Regret}_1^{\mathsf{uPC}}(T) = \mathbb{E}_1\left[\sum_{t=1}^{T}(r_1^* - r_1(P_t))\right] \le \sum_{\ell=0}^{m}\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right].$$

The regret during Phase 0 is equal to

$$\mathbb{E}_1\left[\sum_{t=1}^{\tau_1}(r_1^* - r_1(P_t))\right] = \lceil M_\Phi(P_0^*)\rceil(r_1^* - r_1(P_0^*)).$$

For $1 \le \ell \le m$, the offered price $P_\ell^* \in \{p_1^*, \cdots, p_K^*\}$, so the regret during Phase ℓ is bounded by:

$$\begin{aligned}
&\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}} (r_1^* - r_1(P_t))\right] \\
=&\mathbb{E}_1\left[(\tau_{\ell+1}-\tau_\ell)\times(r_1^* - r_1(P_\ell^*))\right] \\
\le&\mathbb{E}_1\left[\left(M_\Phi(P_\ell^*)e^{(\ell)}+1\right)\times(r_1^* - r_1(P_\ell^*))\right] \\
\le&\left(M_\Phi^* e^{(\ell)}+1\right)\sum_{i=1}^{K}(r_1^* - r_1(p_i^*))\times\mathbb{P}_1[P_\ell^* = p_i^*] \\
\le&\left(M_\Phi^* e^{(\ell)}+1\right)r_1^*\times\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^* = p_i^*].
\end{aligned}$$

By the definition of the uPC policy, the choice of price P_ℓ^* is determined by the sample mean $\bar{X}_{\ell-1}$ in Phase $\ell-1$. Similar to the proof of Theorem 1, letting $\epsilon = \frac{1}{2}\min_{i\neq 1}\left|d_1(P_{\ell-1}^*) - d_i(P_{\ell-1}^*)\right|$, we have

$$\begin{aligned}
&\sum_{i=2}^{K}\mathbb{P}_1[P_\ell^* = p_i^*] \\
\le&\mathbb{P}_1\left[\left|\bar{X}^{\ell-1} - d_1(P_{\ell-1}^*)\right|\ge\epsilon\right] \\
=&\mathbb{E}_1\left[\mathbb{P}_1\left[\left|\bar{X}^{\ell-1} - d_1(P_{\ell-1}^*)\right|\ge\epsilon\middle|P_{\ell-1}^*\right]\right] \\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-(\tau_\ell-\tau_{\ell-1})\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right)\middle|P_{\ell-1}^*\right]\right] \\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-M_\Phi(P_{\ell-1}^*)e^{(\ell-1)}\left(\frac{\epsilon^2}{2\sigma^2}\wedge\frac{\epsilon}{2b}\right)\right)\middle|P_{\ell-1}^*\right]\right] \\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-2e^{(\ell-1)}\right)\middle|P_{\ell-1}^*\right]\right] \\
=&2/(e^{(\ell)})^2.
\end{aligned}$$

So

$$\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}} (r_1^* - r_1(P_t))\right] \le \left(M_\Phi^* e^{(\ell)}+1\right)r_1^*\cdot\frac{2}{(e^{(\ell)})^2} = \frac{2M_\Phi^* r_1^*}{e^{(\ell)}} + \frac{2r_1^*}{(e^{(\ell)})^2}.$$

In sum, the regret of uPC under demand d_1 is bounded by

$$\begin{aligned}
\text{Regret}_1^{\mathsf{uPC}}(T) =& \mathbb{E}_1\left[\sum_{t=1}^{\tau_1}(r_1^* - r_1(P_t))\right] + \sum_{\ell=1}^{m}\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right] \\
\le&(M_\Phi(P_0^*)+1)(r_1^* - r_1(P_0^*)) + \sum_{\ell=1}^{m}\left(\frac{2M_\Phi^* r_1^*}{e^{(\ell)}} + \frac{2r_1^*}{(e^{(\ell)})^2}\right) \\
\le&M_\Phi(P_0^*)(r_1^* - r_1(P_0^*)) + r_1^* + (2M_\Phi^* r_1^* + r_1^*) \\
=&M_\Phi(P_0^*)(r_1^* - r_1(P_0^*)) + 2M_\Phi^* r_1^* + 2r_1^*.
\end{aligned}$$

The minimax regret of demand set Φ is given by

$$\text{Regret}_\Phi^{\mathsf{uPC}}(T) = \max_{i=1,\ldots,K}\text{Regret}_i^{\mathsf{uPC}}(T) \le C_\Phi(P_0^*) + 2M_\Phi^* r^* + 2r^*,$$

where $C_\Phi(P_0^*) = \max_{i\in\{1,\cdots,K\}}\{M_\Phi(P_0^*)(r_i^* - r_i(P_0^*))\}$ and $r^* = \max_{i\in\{1,\cdots,K\}} r_i^*$. $\square$

Appendix EC.3: Proof of Proposition 2

Proof of Proposition 2. Consider a price set $\mathcal{P} = \{1, 2\}$ and two demand functions $d_1(1) = 0.6, d_1(2) = 0.25; d_2(1) = 0.4, d_2(2) = 0.25$. Demand per period has a Bernoulli distribution. It is clear that the optimal prices are $p_1^* = 1, p_2^* = 2$. This demand model violates Assumption 1, because $p_2^* = 2$ is not a discriminative price. We show that for this model, any non-anticipating policy must have a regret of $\Omega(\log T)$.

The one period regret for not using the optimal price is $a = 0.1$ under either demand function. For any policy, we let T_1 be the number of the times that $p = 1$ is used.

We prove the result by contradiction. Suppose $\text{Regret}_2(T) = a \cdot \mathbb{E}_2[T_1] = o(1) \cdot \log T$ and $\text{Regret}_1(T) = a(\mathbb{E}_1[T - T_1]) = o(1) \cdot \log T$. The change-of-measure inequality (see Lemma EC.1) implies that for any event A,

$$\mathbb{P}_2[A] \leq \mathbb{E}_1[1_A \exp(bT_1)].$$

where $b = \log(0.6/0.4)$.

Consider the event: $A = \{T_1 \leq \log T/(2b)\}$, then we have

$$\mathbb{P}_2[A] \leq \mathbb{P}_1[A] \exp(b \cdot \log T/(2b)) = \mathbb{P}_1[A]\sqrt{T}.$$

By Markov's inequality,

$$\mathbb{P}_1[A] = \mathbb{P}_1[T - T_1 \geq T - \log T/(2b)] \leq \frac{\mathbb{E}_1[T - T_1]}{T - \log T/(2b)} = \frac{o(1) \log T}{T - \log T/(2b)}.$$

Thus, we have

$$\mathbb{P}_2[A] \leq \frac{o(1)\sqrt{T} \log T}{T - \log T/(2b)} = o(1).$$

Using Markov's inequality again, we get

$$\mathbb{E}_2[T_1] \geq \frac{\log T}{2b} \mathbb{P}_2[T_1 \geq \frac{\log T}{2b}] = \frac{\log T}{2b}(1 - \mathbb{P}_2[A]) = \frac{\log T}{2b}(1 - o(1)).$$

This contradicts the assumption that $\mathbb{E}_2[T_1] = o(1) \cdot \log T$. $\square$

EC.3.1. Proof of Lemma EC.1

Consider a problem instance (Γ) that satisfies the following conditions:

1. There exists a constant $Q_\Gamma > 0$, such that $\sum_{i=1}^{K} (r_i^* - r_i(p)) \geq Q_\Gamma$ for all $p \in \mathcal{P}$.
2. The demand $D(p) \in \mathbb{N}$ for any price $p \in \mathcal{P}$.
3. Given $p \in \mathcal{P}$, there exists a subset $\mathcal{B}_p \subset \mathbb{N}$, such that for all i, $\mathbb{P}_i[D(p) = d] > 0$ if and only if $d \in \mathcal{B}_p$.
4. There exists a constant $0 < \kappa_\Gamma < 1$, such that $\mathbb{P}_i[D(p) = d]/\mathbb{P}_j[D(p) = d] \geq \kappa_\Gamma$ for all $i, j \in \{1, \ldots, K\}$, $p \in \mathcal{P}, d \in \mathcal{B}_p$.

The first condition states that there is no price $p \in \mathcal{P}$ that simultaneously maximizes the revenue of all demand functions in Φ. This ensures that the problem instance is nontrivial and a learning process is necessary for maximizing the revenue when the demand function is unknown. The second condition is that demand must be integers. The third condition states that all demand functions have the same support for a given price. The fourth condition states that the ratios of probability mass functions of different demand models are bounded. Based on these conditions, we prove the following *change-of-measure lemma.*

LEMMA EC.1. *Let $H_t = (P_1, X_1, \cdots, P_t, X_t)$ be the history observed by the end of period t, and let h_t be a realization of H_t. For any non-anticipating pricing policy π, we have*

$$\mathbb{P}_i^\pi[H_t = h_t] \geq \kappa_\Gamma^t \mathbb{P}_{i'}^\pi[H_t = h_t],$$

for all $i, i' \in \{1, \ldots, K\}$. The constant κ_Γ is defined in the condition (Γ).

Proof of Lemma EC.1. Let $h_t = (p_1, x_1, \ldots, p_t, x_t)$ be a realization of $H_t = (P_1, X_1, \cdots, P_t, X_t)$. We first assume $\mathbb{P}_i^\pi[H_t = h_t] > 0$, so we have

$$\begin{aligned}
\mathbb{P}_i^\pi[H_t = h_t] &= \prod_{s=1}^{t} \mathbb{P}_i^\pi[D(p_s) = x_s] \prod_{s=1}^{t-1} \mathbb{P}_i^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] \\
&= \prod_{s=1}^{t} \left(\mathbb{P}_{i'}^\pi[D(p_s) = x_s] \cdot \frac{\mathbb{P}_i^\pi[D(p_s) = x_s]}{\mathbb{P}_{i'}^\pi[D(p_s) = x_s]} \right) \prod_{s=1}^{t-1} \mathbb{P}_i^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] \qquad \text{(EC.7)} \\
&\geq \prod_{s=1}^{t} \left(\mathbb{P}_{i'}^\pi[D(p_s) = x_s] \cdot \kappa_\Gamma \right) \prod_{s=1}^{t-1} \mathbb{P}_i^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] \qquad \text{(EC.8)} \\
&= \kappa_\Gamma^t \prod_{s=1}^{t} \mathbb{P}_{i'}^\pi[D(p_s) = x_s] \prod_{s=1}^{t-1} \mathbb{P}_i^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] \\
&= \kappa_\Gamma^t \prod_{s=1}^{t} \mathbb{P}_{i'}^\pi[D(p_s) = x_s] \prod_{s=1}^{t-1} \mathbb{P}_{i'}^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] \qquad \text{(EC.9)} \\
&= \kappa_\Gamma^t \mathbb{P}_{i'}^\pi[H_t = h_t].
\end{aligned}$$

Step (EC.7) uses the third condition of (Γ), which states that all demand functions have the same support under a given price, so $\mathbb{P}_{i'}^\pi[D(p_s) = x_s] \neq 0$. Step (EC.8) uses the fourth condition of (Γ). Step (EC.9) holds because price P_{s+1} is determined by policy π and realized history h_s, and is independent of the underlying demand model. Note that if π is a deterministic policy, we always have $\mathbb{P}_i^\pi[P_{s+1} = p_{s+1} \mid H_s = h_s] = 1$ for all i.

Finally, if $\mathbb{P}_i^\pi[H_t = h_t] = 0$, we have $\mathbb{P}_{i'}^\pi[H_t = h_t] = 0$, too. This is again due to the third condition of (Γ), which states that all demand functions have the same support under a given price. $\square$

Appendix EC.4: Proof of Proposition 3

Proof of Proposition 3. Suppose d_1 is the underlying demand function. Let $k \le K-1$ be the number of iterations in the while loop.

The regret under demand d_1 can be composed as

$$\text{Regret}_1^{\mathsf{kPC}}(T) = \mathbb{E}_1\left[\sum_{t=1}^{T}(r_1^* - r_1(P_t))\right] \le \mathbb{E}_1\left[\sum_{\ell=0}^{k}\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right].$$

Let $\epsilon = \frac{1}{2}\min_{i:d_1(P^*_{\ell-1})\ne d_i(P^*_{\ell-1})}\left|d_1(P^*_{\ell-1}) - d_i(P^*_{\ell-1})\right|$. The probability that demand d_1 is eliminated in phase $\ell < k$ is bounded by

$$\begin{aligned}
&\mathbb{P}_1\left[|\bar{X}^{\ell-1} - d_1(P^*_{\ell-1})| \ge |\bar{X}^{\ell-1} - d_i(P^*_{\ell-1})| \text{ for some } i \ne 1\right]\\
\le&\mathbb{P}_1\left[\left|\bar{X}^{\ell-1} - d_1(P^*_{\ell-1})\right| \ge \epsilon\right] \qquad \text{(EC.10)}\\
=&\mathbb{E}_1\left[\mathbb{P}_1\left[\left|\bar{X}^{\ell-1} - d_1(P^*_{\ell-1})\right| \ge \epsilon \middle| P^*_{\ell-1}\right]\right]\\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-(\tau_\ell - \tau_{\ell-1})\left(\frac{\epsilon^2}{2\sigma^2} \wedge \frac{\epsilon}{2b}\right)\right)\middle| P^*_{\ell-1}\right]\right] \qquad \text{(EC.11)}\\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-\tilde{M}_\Phi(P^*_{\ell-1})\log T\left(\frac{\epsilon^2}{2\sigma^2} \wedge \frac{\epsilon}{2b}\right)\right)\middle| P^*_{\ell-1}\right]\right]\\
\le&\mathbb{E}_1\left[2\mathbb{E}_1\left[\exp\left(-\log T\right)\middle| P^*_{\ell-1}\right]\right]\\
=&2/T.
\end{aligned}$$

Inequality (EC.10) is proved in Theorem 1, and (EC.11) uses the Chernoff bound. Since $k \le K-1$, we have

$$\mathbb{P}_1\left[|\bar{X}^{\ell-1} - d_1(P^*_{\ell-1})| \ge |\bar{X}^{\ell-1} - d_i(P^*_{\ell-1})| \text{ for some } i \ne 1, 0 \le \ell < k\right] \le \frac{2(K-1)}{T}.$$

For each of the learning phase ($0 \le \ell \le k-1$), the regret is bounded by

$$\mathbb{E}_1\left[\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right] = \mathbb{E}_1\left[\left\lceil \tilde{M}_A(P^*_\ell)\log T\right\rceil (r_1^* - r_1(P^*_\ell))\right] \le \tilde{M}_\Phi r_1^* \log T + r_1^*.$$

The regret in the earning phase ($\ell = k$) is bounded by

$$\mathbb{E}_1\left[\sum_{t=\tau_k+1}^{T}(r_1^* - r_1(P_t))\right] \le T r_1^* \mathbb{P}_1[P_k \ne p_i^*].$$

So the regret of kPC under demand d_1 is bounded by

$$\begin{aligned}
&\text{Regret}_1^{\mathsf{kPC}}(T) = \mathbb{E}_1\left[\sum_{\ell=0}^{k-1}\sum_{t=\tau_\ell+1}^{\tau_{\ell+1}}(r_1^* - r_1(P_t))\right] + \mathbb{E}_1\left[\sum_{t=\tau_k+1}^{T}(r_1^* - r_1(P_t))\right]\\
\le&(K-1)(\tilde{M}_\Phi r_1^* \log T + r_1^*) + T r_1^* \frac{2(K-1)}{T}\\
=&(K-1)\tilde{M}_\Phi r_1^* \log T + 3(K-1) r_1^*.
\end{aligned}$$

The minimax regret of demand set Φ is given by

$$\text{Regret}_{\Phi}^{\mathsf{kPC}}(T) = \max_{i=1,\ldots,K} \text{Regret}_{i}^{\mathsf{kPC}}(T) \leq (K-1)\tilde{M}_{\Phi} r^* \log T + 3(K-1)r^*.$$

$\square$