

Proofs of Statements

EC.1. Proof of Lemma 1

We first review some notations and concepts necessary to our proof. Let $\mathcal{D} \subseteq \mathbb{R}^{N_1}$ be a convex open set, and $\bar{\mathcal{D}}$ be the closure of \mathcal{D} . We say a function $\psi : \bar{\mathcal{D}} \rightarrow \mathbb{R}$ is *Legendre* if it satisfies the following conditions:

1. ψ is strictly convex and continuously differentiable on \mathcal{D} ;
2. $\lim_{w \rightarrow \bar{\mathcal{D}} \setminus \mathcal{D}} \|\nabla \psi(w)\| = \infty$.

It is easy to verify that our choice of $\psi(w) = \sum_{i=1}^{N_1} -\sqrt{w_i} - \sqrt{1-w_i}$ is Legendre with $\mathcal{D} = (0, 1)^{N_1}$.

Let $\mathcal{D}^* = \nabla \psi(\mathcal{D})$ be the dual space of \mathcal{D} . With our choice of ψ , $\mathcal{D}^* = \mathbb{R}^{N_1}$. For a Legendre function $\psi : \mathcal{D} \rightarrow \mathbb{R}$, its *Legendre-Fenchel transform* (also known as the *convex conjugate*) $\psi^* : \mathcal{D}^* \rightarrow \mathbb{R}$ is defined as

$$\psi^*(u) = \sup_{w \in \bar{\mathcal{D}}} \langle w, u \rangle - \psi(w). \quad (\text{EC.1})$$

The following properties are standard results of convex conjugates. See for example, the reference of (Cesa-Bianchi & Lugosi 2006, Rockafellar 1970), or (Audibert et al. 2014).

Fact 1 *Suppose ψ is Legendre. Then $\psi^{**} = \psi$ and $\nabla \psi^* = (\nabla \psi)^{-1}$. Furthermore, if ψ is also twice continuously differentiable on \mathcal{D} , then $\nabla^2 \psi^*(u) = [\nabla^2 \psi(w)]^{-1}$ for every pair of $w = \nabla \psi^*(u)$ or $u = \nabla \psi(w)$.*

Given a Legendre function $\psi : \bar{\mathcal{D}} \rightarrow \mathbb{R}$, its *Bregman divergence* $D_\psi : \bar{\mathcal{D}} \times \mathcal{D} \rightarrow \mathbb{R}$ is defined as

$$D_\psi(x, y) = \psi(x) - \psi(y) - \langle x - y, \nabla \psi(y) \rangle. \quad (\text{EC.2})$$

If ψ is twice continuously differentiable on \mathcal{D} , then by Taylor expansion with the Lagrangian remainder, we have for every $x, y \in \mathcal{D}$ that

$$D_\psi(x, y) \leq \frac{1}{2} (y - x)^\top \nabla^2 \psi(z) (y - x), \quad (\text{EC.3})$$

where $z = x + \alpha(y - x)$ for some $\alpha \in (0, 1)$.

At time period t , define $\psi_t = \frac{1}{\eta_t} \psi$, where η_t is the step size (or learning rate) at time t . Define function $\phi_t : \mathbb{R}^{N_1} \rightarrow \mathbb{R}$ as

$$\phi_t(u) = \sup_{w \in \Delta^{N_1-1}} \langle w, u \rangle - \psi_t(w). \quad (\text{EC.4})$$

Comparing ϕ_t with the convex conjugate ψ_t^* defined in Eq. (EC.1), the only difference lies in the additional constraint of $w \in \Delta^{N_1-1}$ in the definition of ϕ_t . The following properties are simple and elementary to verify.

Fact 2 *For any Legendre ψ_t let ψ_t^* be its convex conjugate and ϕ_t be defined in Eq. (EC.4). The following properties hold:*

1. *For any $u \in \mathbb{R}^{N_1}$, $\phi_t(u) \leq \psi_t^*(u)$;*
2. *For any $u \in \mathbb{R}^{N_1}$ and $c \in \mathbb{R}$, $\phi_t(u + c\mathbf{1}) = \phi_t(u) + c$, where $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{N_1}$;*
3. *For any $w \in \Delta^{N_1-1}$, $\phi_t(\nabla\psi_t(w)) = \psi_t^*(\nabla\psi_t(w))$;*
4. *For any $u \in \mathbb{R}^{N_1}$, let $w^* = \arg \max_{w \in \Delta^{N_1-1}} \langle w, u \rangle - \psi_t(w)$. Then there exists $\lambda \in \mathbb{R}$ depending on u , such that $\nabla\psi_t(w^*) = u + \lambda\mathbf{1}$.*

Proof of Fact 2. The first property is obvious because ϕ_t has the same objective with ψ_t^* , but with a smaller feasible region. The second property holds because $\phi_t(u + c\mathbf{1}) = \sup_{w \in \Delta^{k-1}} \langle w, u + c\mathbf{1} \rangle - \psi_t(w) = c + \sup_{w \in \Delta^{k-1}} \langle w, u \rangle - \psi_t(w) = c + \phi_t(u)$.

To see the third property, note that $\nabla\psi_t^*(\nabla\psi_t(w)) = w$, thanks to Fact 1. This means that w is the maximizer of $\langle \cdot, \nabla\psi_t(w) \rangle - \psi_t(\cdot)$ on \mathbb{R}^d . Since $w \in \Delta^{N_1-1}$, it is also the maximizer of the same objective on \mathbb{R}^d . Hence, $\phi_t(\nabla\psi_t(w)) = \psi_t^*(\nabla\psi_t(w))$.

For the fourth property, consider the maximization question of $\max_{w \in \mathbb{R}^{N_1-1}} \langle w, u \rangle - \psi_t(w)$ and let w^* be the maximizer. Using the Lagrangian multiplier, we know that $u - \nabla\psi_t(w^*) + \lambda\mathbf{1} = 0$. Hence, $\nabla\psi_t(w^*) = u + \lambda\mathbf{1}$, which is to be demonstrated. \square

Let $e_{i^*} = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^{N_1}$ be the indicator vector corresponding to the price $p(i^*)$. Because $i_t \sim w_t$, we know that $\mathbb{E}[\ell_{t,i_t}] = \mathbb{E}[\langle \ell_t, w_t \rangle]$. Subsequently, the regret $\mathbb{E}[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t}]$ can be decomposed as follows:

$$\begin{aligned} \mathbb{E} \left[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t} \right] &= \mathbb{E} \left[\sum_{t=1}^T \langle \ell_t, e_{i^*} - w_t \rangle \right] \\ &= \underbrace{\mathbb{E} \left[\sum_{t=1}^T \langle \ell_t, -w_t \rangle - \phi_t(\widehat{L}_{t-1}) + \phi_t(\widehat{L}_t) \right]}_{\text{the stability term}} + \underbrace{\mathbb{E} \left[\sum_{t=1}^T \phi_t(\widehat{L}_{t-1}) - \phi_t(\widehat{L}_t) + \langle \ell_t, e_{i^*} \rangle \right]}_{\text{the penalty term}}. \end{aligned} \quad (\text{EC.5})$$

In the rest of this proof, we will upper bound the stability term and the penalty term separately.

EC.1.1. Upper bounding the penalty term

By definition of ϕ_t , $\phi_t(\widehat{L}_{t-1}) = \sup_{w \in \Delta^{N_1-1}} \langle w, \widehat{L}_{t-1} \rangle - \psi_t(w)$ where $\psi_t(w) = \frac{1}{\eta_t} \psi(w)$. Because $w_t \in \Delta^{N_1-1}$ is the maximizer of $\langle w, \widehat{L}_{t-1} \rangle - \psi_t(w)$, we have that

$$\phi_t(\widehat{L}_{t-1}) = \langle w_t, \widehat{L}_{t-1} \rangle - \frac{1}{\eta_t} \psi(w_t). \quad (\text{EC.6})$$

Similarly, for any $w' \in \Delta^{N_1-1}$, it holds that

$$\phi_t(\widehat{L}_t) \geq \langle w', \widehat{L}_t \rangle - \frac{1}{\eta_t} \psi(w'). \quad (\text{EC.7})$$

Combine Eqs. (EC.6, EC.7) and set w' in Eq. (EC.7) as $w' = w_{t+1}$ for $t < T$, and $w' = e_{i^*}$ for $t = T$.

We then have

$$\begin{aligned} & \sum_{t=1}^T \phi_t(\widehat{L}_{t-1}) - \phi_t(\widehat{L}_t) \\ &= \left[\sum_{t=1}^T \langle w_t, \widehat{L}_{t-1} \rangle - \frac{1}{\eta_t} \psi(w_t) \right] - \left[\sum_{t=1}^{T-1} \langle w_{t+1}, \widehat{L}_t \rangle - \frac{1}{\eta_t} \psi(w_{t+1}) \right] - \langle \widehat{L}_T, e_{i^*} \rangle + \frac{1}{\eta_T} \psi(e_{i^*}) \\ &= \left[\sum_{t=2}^T \left(\frac{1}{\eta_{t-1}} - \frac{1}{\eta_t} \right) \psi(w_t) \right] - \frac{1}{\eta_1} \psi(w_1) - \langle \widehat{L}_T, e_{i^*} \rangle + \frac{1}{\eta_T} \psi(e_{i^*}), \end{aligned}$$

where the last equality holds because $\widehat{L}_0 = 0$ by definition. Notice that $\widehat{L}_T = \sum_{t=1}^T \widehat{\ell}_t$ satisfies $\mathbb{E}[\widehat{L}_T] = \sum_{t=1}^T \mathbb{E}[\widehat{\ell}_t] = \sum_{t=1}^T \ell_t$. Hence,

$$\begin{aligned} & \mathbb{E} \left[\sum_{t=1}^T \phi_t(\widehat{L}_{t-1}) - \phi_t(\widehat{L}_t) + \langle \ell_t, e_{i^*} \rangle \right] = \mathbb{E} \left[\sum_{t=2}^T \left(\frac{1}{\eta_{t-1}} - \frac{1}{\eta_t} \right) \psi(w_t) - \frac{1}{\eta_1} \psi(w_1) + \frac{1}{\eta_T} \psi(e_{i^*}) \right] \\ &= \mathbb{E} \left[\sum_{t=2}^T \left(\frac{1}{\eta_{t-1}} - \frac{1}{\eta_t} \right) (\psi(w_t) - \psi(e_{i^*})) - \frac{1}{\eta_1} (\psi(w_1) - \psi(e_{i^*})) \right], \quad (\text{EC.8}) \end{aligned}$$

where the last equality holds because the terms involving $\psi(e_{i^*})$ sum to $\frac{1}{\eta_T} \psi(e_{i^*})$.

Next, we analyze the differences between $\psi(w_t)$ and $\psi(e_{i^*})$. Recall that, for $w \in \Delta^{N_1-1}$, $\psi(w)$ is defined as $\psi(w) = \sum_{i=1}^{N_1} -\sqrt{w_i} - \sqrt{1-w_i}$. Also, because each component of e_{i^*} is either 1 or 0, we have $\psi(e_{i^*}) = -N_1$. Subsequently,

$$\psi(e_{i^*}) - \psi(w_t) = \sum_{i=1}^{N_1} \sqrt{w_{ti}} + \sqrt{1-w_{ti}} - 1 \leq \min\{\sqrt{w_{ti}}, \sqrt{1-w_{ti}}\}. \quad (\text{EC.9})$$

Plugging Eq. (EC.9) into Eq. (EC.8), and noting that $\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} = (\sqrt{t} - \sqrt{t-1})/\eta_0 \leq 1/(\eta_0\sqrt{t})$, we have that

$$\begin{aligned}
\mathbb{E} \left[\sum_{t=1}^T \phi_t(\widehat{L}_{t-1}) - \phi_t(\widehat{L}_t) + \langle \ell_t, e_{i^*} \rangle \right] &\leq \mathbb{E} \left[\sum_{t=1}^T \sum_{i=1}^{N_1} \frac{\min\{\sqrt{w_{ti}}, \sqrt{1-w_{ti}}\}}{\eta_0\sqrt{t}} \right] \\
&\leq \frac{1}{\eta_0} \mathbb{E} \left[\sum_{t=1}^T \sqrt{\frac{1-w_{ti^*}}{t}} + \sum_{i \neq i^*} \sqrt{\frac{w_{ti}}{t}} \right] = \frac{1}{\eta_0} \mathbb{E} \left[\sum_{t=1}^T \sqrt{\frac{\sum_{i \neq i^*} w_{ti}}{t}} + \sum_{i \neq i^*} \sqrt{\frac{w_{ti}}{t}} \right] \\
&\leq \frac{2}{\eta_0} \mathbb{E} \left[\sum_{t=1}^T \sum_{i \neq i^*} \sqrt{\frac{w_{ti}}{t}} \right]. \tag{EC.10}
\end{aligned}$$

EC.1.2. Upper bounding the stability term

Recall the definition that $w_t = \arg \max_{w \in \Delta^{N_1-1}} \langle w, \widehat{L}_{t-1} \rangle - \psi_t(w)$. By the fourth property of Fact 2, there exists $\lambda_t \in \mathbb{R}$ such that $\nabla \psi_t(w_t) = \widehat{L}_{t-1} + \lambda_t \mathbf{1}$. Subsequently,

$$\begin{aligned}
-\phi_t(\widehat{L}_{t-1}) + \phi_t(\widehat{L}_t) &= -\phi_t(\nabla \psi_t(w_t) - \lambda_t \mathbf{1}) + \phi_t(\nabla \psi_t(w_t) - \lambda_t \mathbf{1} + \widehat{\ell}_t) \\
&= -\phi_t(\nabla \psi_t(w_t)) + \phi_t(\nabla \psi_t(w_t) + \widehat{\ell}_t), \tag{EC.11}
\end{aligned}$$

where the second equality holds thanks to the second property of Fact 2 (so that the $\lambda_t \mathbf{1}$ terms are canceled), and that $\widehat{L}_t = \widehat{L}_{t-1} + \widehat{\ell}_t$. By the first and the third properties of Fact 2, we know that $\phi_t(\nabla \psi_t(w_t)) = \psi_t^*(\nabla \psi_t(w_t))$ and $\phi_t(\nabla \psi_t(w_t) + \widehat{\ell}_t) \leq \psi_t^*(\nabla \psi_t(w_t) + \widehat{\ell}_t)$. Subsequently,

$$\begin{aligned}
-\langle \ell_t, w_t \rangle - \phi_t(\widehat{L}_{t-1}) + \phi_t(\widehat{L}_t) &\leq -\langle \ell_t, w_t \rangle - \psi_t^*(\nabla \psi_t(w_t)) + \psi_t^*(\nabla \psi_t(w_t) + \widehat{\ell}_t) \\
&= D_{\psi_t^*}(\nabla \psi_t(w_t) + \widehat{\ell}_t, \nabla \psi_t(w_t)), \tag{EC.12}
\end{aligned}$$

where the last equality holds because $\nabla \psi_t^*(\nabla \psi_t(w_t)) = w_t$, thanks to the $\nabla \psi_t^* = (\nabla \psi_t)^{-1}$ property in Fact 1. Using Eq. (EC.3) and the relationship between $\nabla_{\psi_t^*}^2, \nabla_{\psi_t}^2$ in Fact 1, we have that

$$\begin{aligned}
D_{\psi_t^*}(\nabla \psi_t(w_t) + \widehat{\ell}_t, \nabla \psi_t(w_t)) &\leq \frac{1}{2} \widehat{\ell}_t^\top \nabla^2 \psi_t^*(\nabla \psi_t(w_t) + \alpha_t \widehat{\ell}_t) \widehat{\ell}_t \\
&= \frac{1}{2} \widehat{\ell}_t^\top [\nabla^2 \psi_t(\nabla \psi_t^*(\nabla \psi_t(w_t) + \alpha_t \widehat{\ell}_t))]^{-1} \widehat{\ell}_t, \tag{EC.13}
\end{aligned}$$

where $\alpha_t \in (0, 1)$ is a certain interpolation parameter.

The following lemma upper bounds the discrepancy between $\nabla \psi_t^*(\nabla \psi_t(w_t) + \alpha_t \widehat{\ell}_t)$ and $\nabla \psi_t^*(\nabla \psi_t(w_t)) = w_t$.

LEMMA EC.1. *Let $\tilde{w} = \nabla \psi_t^*(\nabla \psi_t(w) + \alpha \delta)$ for some $w \in \Delta^{K_1-1}$, $\alpha \in (0, 1)$ and $\delta \in \mathbb{R}^{N_1-1}$ satisfying $-\frac{\bar{p}}{w_i} \leq \delta_i \leq \bar{p}$. The potential function is chosen as $\psi_t(w) = \frac{1}{\eta_t} \psi(w)$ where $\psi(w) = \sum_{i=1}^{N_1} -\sqrt{w_i} - \sqrt{1-w_i}$. Suppose $\eta_t \leq \frac{1}{4\bar{p}}(1 - \frac{1}{\sqrt{2}})$. Then it holds for all $i \in [N_1]$ that $2w_i - 1 \leq \tilde{w}_i \leq 2w_i$.*

We will prove Lemma EC.1 in the next section. For the rest of the proof, note that the condition $\eta_0 = 0.07/\bar{p} \leq \frac{1}{4\bar{p}}(1 - \frac{1}{\sqrt{2}})$ in Lemma 1 implies the condition on η_t in Lemma EC.1 holds, because $\eta_t = \eta_0/\sqrt{t} \leq \eta_0$. Note also that, for any $w \in [0, 1]^{N_1}$ and $j \in [N_1]$,

$$\partial_{jj}^2 \psi(w) = \frac{1}{4} \left(\frac{1}{w_j^{3/2}} + \frac{1}{(1-w_j)^{3/2}} \right).$$

With Lemma EC.1 and the notation that $\tilde{w}_t = w_t + \alpha_t \hat{\ell}_t$, and conditioned on the event that $i_t = i$, it holds that

$$\begin{aligned} \hat{\ell}_t^\top [\nabla^2 \psi_t(\tilde{w}_t)]^{-1} \hat{\ell}_t &= \sum_{j \neq i} \bar{p}^2 [\partial_{jj}^2 \psi_t(\tilde{w}_{tj})]^{-1} + \left(\frac{p(i)d_t - \bar{p}}{w_{ti}} + \bar{p} \right)^2 [\partial_{ii}^2 \psi_t(\tilde{w}_{ti})]^{-1} \\ &\leq \sum_{j=1}^{N_1} \bar{p}^2 [\partial_{jj}^2 \psi_t(\tilde{w}_{tj})]^{-1} + \frac{\bar{p}^2}{w_{ti}^2} [\partial_{ii}^2 \psi_t(\tilde{w}_{ti})]^{-1} \end{aligned} \quad (\text{EC.14})$$

$$\begin{aligned} &= 4\bar{p}^2 \eta_t \sum_{j=1}^{N_1} \left(\frac{1}{\tilde{w}_{tj}^{3/2}} + \frac{1}{(1-\tilde{w}_{tj})^{3/2}} \right)^{-1} + \frac{4\bar{p}^2 \eta_t}{w_{ti}^2} \left(\frac{1}{\tilde{w}_{ti}^{3/2}} + \frac{1}{(1-\tilde{w}_{ti})^{3/2}} \right)^{-1} \\ &\leq 8\sqrt{2}\bar{p}^2 \eta_t \sum_{j=1}^{N_1} \min \{ w_{tj}^{3/2}, (1-w_{tj})^{3/2} \} + \frac{8\sqrt{2}\bar{p}^2 \eta_t}{w_{ti}^2} \min \{ w_{ti}^{3/2}, (1-w_{ti})^{3/2} \}. \end{aligned} \quad (\text{EC.15})$$

Here, Eq. (EC.14) holds because $0 \leq p(i)d_t \leq \bar{p}$ and hence $|\frac{p(i)d_t - \bar{p}}{w_{ti}} + \bar{p}| \leq \frac{\bar{p}}{w_{ti}}$. Eq. (EC.15) holds because $(\tilde{w}_{tj}^{-3/2} + (1-\tilde{w}_{tj})^{-3/2})^{-1} \leq \min \{ \tilde{w}_{tj}^{3/2}, (1-\tilde{w}_{tj})^{3/2} \} \leq \min \{ (2w_{tj})^{3/2}, (2-2w_{tj})^{3/2} \}$, thanks to Lemma EC.1. Because $i_t = i$ with probability w_{ti} , the right-hand side of Eq. (EC.15) can be further upper bounded by

$$\begin{aligned} &8\sqrt{2}\bar{p}^2 \eta_t \sum_{j=1}^{N_1} \min \{ w_{tj}^{3/2}, (1-w_{tj})^{3/2} \} + 8\sqrt{2}\bar{p}^2 \eta_t \sum_{i=1}^{N_1} \frac{\min \{ w_{ti}^{3/2}, (1-w_{ti})^{3/2} \}}{w_{ti}} \\ &\leq 16\sqrt{2}\bar{p}^2 \eta_t \sum_{i=1}^{N_1} \frac{\min \{ w_{ti}^{3/2}, (1-w_{ti})^{3/2} \}}{w_{ti}} \leq 16\sqrt{2}\bar{p}^2 \eta_t \sum_{i=1}^{N_1} \min \{ \sqrt{w_{ti}}, \sqrt{1-w_{ti}} \} \\ &\leq 16\sqrt{2}\bar{p}^2 \eta_t \left[\sqrt{1-w_{ti^*}} + \sum_{i \neq i^*} \sqrt{w_{ti}} \right] = 16\sqrt{2}\bar{p}^2 \eta_t \left[\sqrt{\sum_{i \neq i^*} w_{ti}} + \sum_{i \neq i^*} \sqrt{w_{ti}} \right] \\ &\leq 32\sqrt{2}\bar{p}^2 \eta_t \left[\sum_{i \neq i^*} \sqrt{w_{ti}} \right] = 32\sqrt{2}\bar{p}^2 \eta_0 \left[\sum_{i \neq i^*} \sqrt{\frac{w_{ti}}{t}} \right]. \end{aligned} \quad (\text{EC.16})$$

EC.1.3. Proof of Lemma EC.1

Because $\nabla \psi_t^* = (\nabla \psi_t)^{-1}$ thanks to Fact 1, we have that

$$\nabla \psi_t(\tilde{w}) = \nabla \psi_t(w) + \alpha \delta.$$

Fix an arbitrary $i \in [N_1]$ and let w_i, \tilde{w}_i be the i th components of w and \tilde{w} , respectively. The above inequality then implies that

$$\partial_i \psi_t(w_i) - \alpha \delta_i \leq \partial_i \psi_t(\tilde{w}_i) \leq \partial_i \psi_t(w_i) + \alpha \delta_i. \quad (\text{EC.17})$$

We first prove $\tilde{w}_i \leq 2w_i$. If $w_i \geq 1/2$ then the inequality automatically holds because $\tilde{w} \in \text{Range}(\nabla \psi_t^*) \subseteq [0, 1]^{N_1}$. Hence, we shall assume that $w_i < 1/2$. Assume by way of contradiction that $\tilde{w}_i > 2w_i$. Because $\partial_i \psi_t(w_i) = \frac{1}{\eta_t} \left[-\frac{1}{2\sqrt{w_i}} + \frac{1}{2\sqrt{1-w_i}} \right]$ is strictly monotonically increasing with w_i , we have that

$$\alpha \delta_i = \partial_i \psi_t(\tilde{w}_i) - \partial_i \psi_t(w_i) > \partial_i \psi_t(2w_i) - \partial_i \psi_t(w_i).$$

Subsequently,

$$\alpha \delta_i > \frac{1}{\eta_t} \left[-\frac{1}{2\sqrt{2w_i}} + \frac{1}{2\sqrt{1-2w_i}} + \frac{1}{2\sqrt{w_i}} - \frac{1}{2\sqrt{1-w_i}} \right] \geq \frac{1}{2} \left(1 - \frac{1}{\sqrt{2}} \right) \frac{1}{\eta_t w_i} \geq \bar{p},$$

where the last inequality holds because $w_i \leq 1$ and $\eta_t \leq \frac{1}{2\bar{p}}(1 - \frac{1}{\sqrt{2}})$. This contradicts the condition that $\alpha \in (0, 1)$ and $\delta_i \leq \bar{p}$.

We next prove $\tilde{w}_i \geq 2w_i - 1$. If $w_i \leq 1/2$ then the inequality automatically holds because $\tilde{w} \in \text{Range}(\nabla \psi_t^*) \subseteq [0, 1]^{N_1}$. Hence we shall assume that $w_i > 1/2$. Assume by way of contradiction that $\tilde{w}_i < 2w_i - 1$. Again by the strict monotonicity of $\partial_i \psi_t$, we have that

$$\alpha \delta_i = \partial_i \psi_t(\tilde{w}_i) - \partial_i \psi_t(w_i) < \partial_i \psi_t(2w_i - 1) - \partial_i \psi_t(w_i).$$

Subsequently,

$$\begin{aligned} \alpha \delta_i &< \frac{1}{\eta_t} \left[-\frac{1}{2\sqrt{2w_i-1}} + \frac{1}{2\sqrt{2(1-w_i)}} + \frac{1}{2\sqrt{w_i}} - \frac{1}{2\sqrt{1-w_i}} \right] \leq \frac{1}{\eta_t} \left[\frac{1}{2\sqrt{2(1-w_i)}} - \frac{1}{2\sqrt{1-w_i}} \right] \\ &\leq -\frac{1}{2} \left(1 - \frac{1}{\sqrt{2}} \right) \frac{1}{\eta_t \sqrt{1-w_i}} \leq -\frac{1}{2} \left(1 - \frac{1}{\sqrt{2}} \right) \frac{1}{2\eta_t w_i}, \end{aligned}$$

where the last inequality holds because $\sqrt{1-w_i} \leq 2w_i$ for all $w_i \in (1/2, 1]$. Because $\eta_t \leq \frac{1}{4\bar{p}}(1 - \frac{1}{\sqrt{2}})$, we have that

$$\alpha \delta_i < -\bar{p}/w_i,$$

which contradicts the condition that $\alpha \in (0, 1)$ and $\delta_i \geq -\bar{p}/w_i$.

EC.1.4. Putting everything together

Combine Eqs. (EC.5, EC.10, EC.16). We have that

$$\mathbb{E} \left[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t} \right] \leq \left(32\sqrt{2\bar{p}^2}\eta_0 + \frac{2}{\eta_0} \right) \times \mathbb{E} \left[\sum_{t=1}^T \sum_{i \neq i^*} \sqrt{\frac{w_{ti}}{t}} \right].$$

Plugging in the scaling that $\eta_0 = 0.07/\bar{p}$ we complete the proof of Lemma 1.

EC.2. Proof of Theorem 1

Recall that $\ell_{t,i} = p(i)f_t(p(i))$ and $i^* = \arg \max_{i \in [N_1]} p(i)f_0(p(i))$ is the revenue maximizer among prices $\{p(i)\}_{i=1}^{N_1}$ for typical customers.

We then have

$$\mathfrak{R}_{T,\varepsilon T}(\text{Alg.3}; f_0, x_0) \leq T |p^\circ f_0(p^\circ) - p(i^*)f_0(p(i^*))| + \mathbb{E} \left[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t} \right]. \quad (\text{EC.18})$$

To upper bound the second term on the right-hand side of Eq. (EC.18), note that because $\{p(i)\}_{i=1}^{N_1}$ are N_1 prices evenly partitioning $[\underline{p}, \bar{p}]$, there exists $i^\# \in [N_1]$ such that $|p(i^\#) - p^\circ| \leq (\bar{p} - \underline{p})/(2N_1)$. Other the other hand, because $r(d) = df_0^{-1}(d)$ is strongly smooth thanks to Assumption (A2), we have that $r(d^\circ) - r(d(i^\#)) \leq \frac{M^2}{2}|d^\circ - d(i^\#)|^2$, where $d^\circ = f_0(p^\circ)$ and $d(i^\#) = f_0(p(i^\#))$. Subsequently,

$$\begin{aligned} p^\circ f_0(p^\circ) - p(i^*)f_0(p(i^*)) &\leq p^\circ f_0(p^\circ) - p(i^\#)f_0(p(i^\#)) = d^\circ f_0^{-1}(d^\circ) - d(i^\#)f_0^{-1}(d(i^\#)) \\ &= r(d^\circ) - r(d(i^\#)) \\ &\leq \frac{M^2}{2}|d^\circ - d(i^\#)|^2 \leq \frac{M^2}{2}\bar{L}_d^2|p^\circ - p(i^\#)|^2 \end{aligned} \quad (\text{EC.19})$$

$$\leq \frac{M^2\bar{L}_d^2}{2} \frac{(\bar{p} - \underline{p})^2}{4N_1^2} = \frac{M^2\bar{L}_d^2(\bar{p} - \underline{p})^2}{8\sqrt{T}}. \quad (\text{EC.20})$$

Here the second inequality in Eq. (EC.19) holds because $|d^\circ - d(i^\#)| = |f_0^{-1}(p^\circ) - f_0^{-1}(p(i^\#))| \leq \bar{L}_d|p^\circ - p(i^\#)|$. Subsequently,

$$T |p^\circ f_0(p^\circ) - p(i^*)f_0(p(i^*))| \leq \frac{1}{8}M^2\bar{L}_d^2(\bar{p} - \underline{p})^2\sqrt{T}. \quad (\text{EC.21})$$

We next turn to the third term on the right-hand side of Eq. (EC.18). By Lemma 2, for every $\mathcal{I} \subseteq [N_1] \setminus \{i^*\}$, it holds that

$$\mathbb{E} \left[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t} \right] \leq 64\bar{p}\sqrt{|\mathcal{I}|T} + \sum_{t=1}^T \sum_{i \notin \mathcal{I}, i \neq i^*} \frac{1}{2} \frac{(32\bar{p})^2}{t\Delta\mu_i} + \frac{1}{2}\bar{p}\varepsilon T, \quad (\text{EC.22})$$

where $\Delta\mu_i = p(i^*)f_0(p(i^*)) - p(i)f_0(p(i))$. For any $i \in [N_1]$, define $\gamma = |i - i^\#|$ where $i^\# = \arg \min_{i \in [N_1]} |p(i) - p^\circ|$. By Lemma 3, if $\gamma \geq \frac{1}{2} + \frac{M\bar{L}_p}{\sqrt{2}\sigma\bar{L}_p}$ then it holds that

$$\Delta\mu_i \geq p(i^\#)f_0(p(i^\#)) - p(i)f_0(p(i)) \geq \frac{\sigma^2\bar{L}_p^2(\gamma - 1/2)^2\zeta^2}{4},$$

where $\zeta = (\bar{p} - p^\circ)/N_1$. Now let $\mathcal{I} = \{i \in [N_1] : i \neq i^*, |i - i^\#| \leq \frac{1}{2} + \frac{M\bar{L}_p}{\sqrt{2}\sigma\bar{L}_p}\}$. Clearly $|\mathcal{I}| \leq 2 + \sqrt{2}M\bar{L}_p/(\sigma\bar{L}_p)$. Subsequently, Eq. (EC.22) can be reduced to

$$\begin{aligned} & \mathbb{E} \left[\sum_{t=1}^T \ell_{t,i^*} - \ell_{t,i_t} \right] \\ & \leq 64\bar{p}\sqrt{(2 + \sqrt{2}M\bar{L}_p/(\sigma\bar{L}_p))T} + \frac{1}{2}\bar{p}\varepsilon T + \sum_{\gamma=1}^{N_1} 512\bar{p}^2(\ln T + 1) \times \left[\frac{\sigma^2\bar{L}_p^2(\gamma - 1/2)^2\zeta^2}{4} \right]^{-1} \\ & \leq 128\sqrt{\frac{M\bar{L}_p T}{\sigma\bar{L}_p}} + \frac{1}{2}\bar{p}\varepsilon T + \frac{2048\bar{p}^2 \ln(eT)}{\sigma^2\bar{L}_p^2\zeta^2} \times \sum_{\gamma=1}^{N_1} \frac{1}{(\gamma - 1/2)^2} \\ & \leq 128\sqrt{\frac{M\bar{L}_p T}{\sigma\bar{L}_p}} + \frac{1}{2}\bar{p}\varepsilon T + \frac{2048\bar{p}^2 \ln(eT)}{\sigma^2\bar{L}_p^2\zeta^2} \times \frac{\pi}{3}. \end{aligned} \quad (\text{EC.23})$$

Note that $\zeta = (\bar{p} - p^\circ)/N_1$ and $N_1 \approx T^{1/4}$. Combining Eqs. (EC.18, EC.21, EC.23) we obtain

$$\begin{aligned} \mathfrak{R}_{T,\varepsilon T}(\text{Alg. 3}; f_0, x_0) & \leq \frac{1}{8}M^2\bar{L}_d^2(\bar{p} - p^\circ)^2\sqrt{T} + 128\sqrt{\frac{M\bar{L}_p T}{\sigma\bar{L}_p}} + \frac{1}{2}\bar{p}\varepsilon T + \frac{2145\bar{p}^2\sqrt{T}\ln(eT)}{\sigma^2\bar{L}_p^2(\bar{p} - p^\circ)^2} \\ & = \tilde{O}(\sqrt{T} + \varepsilon T), \end{aligned}$$

which is to be demonstrated.

EC.3. Proof of Proposition 1.

For notation simplicity we shall omit the π superscript in this proof. For every $t < T$, it holds that $\mathbb{E}[\sum_{t' \leq t} d_{t'}] \leq \min\{x_T, x_0 t + \mathbb{E}[\sum_{t' \leq t} \delta_{t'}] + E\} \leq x_0 t + \mathbb{E}[\bar{\delta}] + E$. Using Hoeffding's inequality and the union bound, we have with probability $1 - T^{-1}$ that

$$\frac{1}{t} \sum_{t' \leq t} d_{t'} \leq x_0 + \frac{\mathbb{E}[\bar{\delta}] + E}{t} + \sqrt{\frac{\ln T}{t}}$$

for every $t < T$. Or equivalently,

$$\sum_{t' \leq t} d_{t'} \leq x_0 t + \mathbb{E}[\bar{\delta}] + E + \sqrt{t \ln T}. \quad (\text{EC.24})$$

Now let $T^+ \leq T$ be the random variable of the last episode such that $x(T^+) > 0$. This implies that $\sum_{t' \leq T^+} d_{t'} > x_T = x_0 T$. Comparing this with Eq. (EC.24), we have that

$$\mathbb{E}[x_0 T^+ + \bar{\delta} + \varepsilon T + \sqrt{T^+ \log T}] \geq x_0 T.$$

Re-arranging the inequality and noting that $T^+ \leq T$, $x_0 \in (0, 1]$, we have that

$$\mathbb{E}[T - T^+] \leq \frac{1}{x_0} \left(\mathbb{E}[\bar{\delta}] + \varepsilon T + \sqrt{T \log T} \right).$$

This completes the proof, because $\bar{p}(T - T^+)$ is an upper bound on the regret incurred by lost sales over the T time periods.

To prove the high-probability claim, note that the above argument remains valid if one condition on the event that $\bar{\delta} \leq \bar{B}$ and taking expectations over the randomness of $\{f_t(p_t) - d_t\}$ only. \square

EC.4. Proofs of Technical Results in Sec. 5.1

To prove the Theorem 2, we first introduce some technical lemmas. The following lemma shows that for an epoch τ , $[\underline{d}(\tau), \bar{d}(\tau)]$ covers the expected demand at the price $p(\tau)$ with high probability.

LEMMA EC.2. *Suppose for some epoch τ , the inventory levels kept positive throughout the epoch. Then with probability $1 - O(T^{-2})$, $\underline{d}(\tau) \leq f_0(p(\tau)) \leq \bar{d}(\tau)$.*

Proof of Lemma EC.2. To prove Lemma EC.2 we first present and prove another lemma, where we try to bound the gap between “the averaged realized demand over T' periods” during which the price p was offered and the actual demand at price p .

LEMMA EC.3. *Let \mathcal{T} be a set of T' selling periods during which the inventory level remained positive. Let $p \in [p, \bar{p}]$ be a fixed price, which is offered at each selling period $t \in \mathcal{T}$ with probability $q \in (0, 1]$. Let \hat{d} be the total realized demands over $t \in \mathcal{T}$ during which price p is offered. Suppose also that the total number of corrupted periods during the T' selling periods considered is upper bounded by εT almost surely. Then for any $\delta \in (0, 1)$, with probability $1 - \delta$ it holds that*

$$\left| \frac{\hat{d}}{qT'} - f_0(p) \right| \leq \min\{1, \varepsilon T/T'\} + \sqrt{\frac{\log(2/\delta)}{qT'}} + \frac{2 \log(2/\delta)}{qT'}.$$

Proof of Lemma EC.3. For each $t \in \mathcal{T}$, let $\nu_t = \mathbf{1}\{p_t = p\}$ be the indicator random variable denoting whether price p is offered at time t . Recall that $\iota_t \in \{0, 1\}$ indicates whether t is an outlier period. Define also $z_t := f_t(p)$. We then have

$$\widehat{d} = \sum_{t \in \mathcal{T}} \nu_t \iota_t z_t + \nu_t (1 - \iota_t) f_0(p) = \sum_{t \in \mathcal{T}} \nu_t [\iota_t (z_t - f_0(p)) + f_0(p)].$$

Because ν_t is statistically independent from both ι_t and z_t , and that $\Pr[\nu_t = 1] = q$, we have that

$$\mathbb{E}[\widehat{d}] = qZ + qT' f_0(p),$$

where $Z = \sum_{t \in \mathcal{T}} \iota_t z_t$.

To upper bound the deviation of \widehat{d} from $\mathbb{E}[\widehat{d}]$, we need the following result, which is Theorem 1.2A cited from (Victor 1999).

LEMMA EC.4. *Let $\{w_i, \mathcal{F}_i\}$ be a martingale difference sequence with $\mathbb{E}[w_j | \mathcal{F}_{j-1}] = 0$, $\mathbb{E}[w_j^2 | \mathcal{F}_{j-1}] = \sigma_j^2$, $V_n^2 = \sum_{j=1}^n \sigma_j^2$. Furthermore, assume that $\Pr[|w_j| \leq c | \mathcal{F}_{j-1}] = 1$ for some $0 < c < \infty$. Then, for all $\epsilon, y > 0$, it holds that*

$$\Pr \left[\sum_{i=1}^n w_i \geq \epsilon, V_n^2 \leq y \text{ for some } n \right] \leq \exp \left\{ -\frac{\epsilon^2}{2(y + c\epsilon)} \right\}.$$

We now go back to the definition of \widehat{d} and write it as $\widehat{d} - \mathbb{E}[\widehat{d}] = \sum_{t \in \mathcal{T}} w_t - \mathbb{E}[w_t | \mathcal{F}_{t-1}]$, where $w_t = \nu_t [\iota_t (z_t - f_0(p)) + f_0(p)]$ and \mathcal{F}_{t-1} denotes the filtration prior to time t . Clearly $\{w_t - \mathbb{E}[w_t | \mathcal{F}_{t-1}]\}_{t \in \mathcal{T}}$ forms a martingale difference sequence with zero mean. Furthermore, $|w_t| \leq 1$ almost surely, and $\mathbb{E}[w_t^2 | \mathcal{F}_{t-1}] \leq \mathbb{E}[\nu_t | \mathcal{F}_{t-1}] = q$. This means that V_n^2 in Lemma EC.4 satisfies $V_n^2 \leq T'q$ almost surely. With $n = T'$, $c = 1$, $y = T'q$ and ϵ appropriately set in Lemma EC.4, we have with probability $1 - \delta$ that

$$\left| \sum_{t \in \mathcal{T}} w_t - \mathbb{E}[w_t | \mathcal{F}_{t-1}] \right| \leq \log(2/\delta) + \sqrt{\log^2(2/\delta) + y \log(2/\delta)} \leq 2 \log(2/\delta) + \sqrt{T'q \log(2/\delta)}.$$

Recall that $\sum_{t \in \mathcal{T}} w_t = \widehat{d}$ and $\sum_{t \in \mathcal{T}} \mathbb{E}[w_t | \mathcal{F}_{t-1}] = \mathbb{E}[\widehat{d}] = qZ + qT' f_0(p)$. We then have with probability $1 - \delta$ that

$$|\widehat{d} - qT' f_0(p)| \leq qZ + 2 \log(2/\delta) + \sqrt{T'q \log(2/\delta)} \leq q \max\{T', \varepsilon T\} + 2 \log(2/\delta) + \sqrt{T'q \log(2/\delta)},$$

where the last inequality holds since at most εT periods in \mathcal{T} can be corrupted. Dividing both sides of the above inequality by qT' we proved Lemma EC.3. \square

Invoking Lemma EC.3 with $T' = T(\tau)$, $q = 1$ and $\delta = 1/T^2$. We then have with probability $1 - O(T^{-2})$ that

$$\left| \frac{d}{T(\tau)} - f_0(p(\tau)) \right| \leq \min \left\{ 1, \frac{\varepsilon T}{T(\tau)} \right\} + \sqrt{\frac{\log(2T^2)}{T(\tau)}} + \frac{2 \log(2T^2)}{T(\tau)} = C_\varepsilon(\tau),$$

which proves Lemma EC.2. \square

Given Lemma EC.2, we are ready to state the next key Lemma for proving Theorem 2. The next lemma provides an upper bound on the length of the searching interval $|I(\tau)| := (b(\tau) - a(\tau))$.

LEMMA EC.5. *With probability $1 - O(T^{-2} \log T)$ the following holds: at the beginning of every epoch τ , $p^c \in I(\tau)$ and the length of $I(\tau)$ is upper bounded by:*

$$|I(\tau)| = (b(\tau) - a(\tau)) \leq 2\bar{L}_d C_\varepsilon(\tau - 1).$$

Proof of Lemma EC.5. Throughout this proof we will assume that $f_0(p(\tau)) \in [\underline{d}(\tau), \bar{d}(\tau)]$ for every τ , at the end of each epoch. By Lemma EC.2, this occurs with probability $1 - O(T^{-2} \log T)$.

We first prove $p^c \in I(\tau)$ for any epoch τ . Recall that p^c is the unique price for which $f_0(p^c) = x_0$. Also note that f_0 is strictly monotonically decreasing in p . Hence, in the cases of $x_0 \notin [\underline{d}(\tau), \bar{d}(\tau)]$, it is clear that p^c remains in the shrunk interval $I(\tau + 1)$. For the other case of $x_0 \in [\underline{d}(\tau), \bar{d}(\tau)]$, denote $d(\tau) := f_0(p(\tau))$. Since f_0 is strictly decreasing, f_0^{-1} is also a strictly decreasing function. By Assumption (A1), we have that

$$p^c = f_0^{-1}(x_0) \geq f_0^{-1}(\bar{d}(\tau)) \geq f_0^{-1}(d(\tau)) - \bar{L}_d(\bar{d}(\tau) - d(\tau)) \geq f_0^{-1}(d(\tau)) - \bar{L}_d(\bar{d}(\tau) - \underline{d}(\tau));$$

$$p^c = f_0^{-1}(x_0) \leq f_0^{-1}(\underline{d}(\tau)) \leq f_0^{-1}(d(\tau)) + \bar{L}_d(d(\tau) - \underline{d}(\tau)) \leq f_0^{-1}(d(\tau)) + \bar{L}_d(\bar{d}(\tau) - \underline{d}(\tau)).$$

Note that $f_0^{-1}(d(\tau)) = p(\tau)$ and $\bar{d}(\tau) - \underline{d}(\tau) = 2C_\varepsilon(\tau)$. This justifies that $p^c \in I(\tau + 1)$ in the case of $x_0 \in [\underline{d}(\tau), \bar{d}(\tau)]$.

We next use induction to prove that $|I(\tau)| \leq 2\bar{L}_d C_\varepsilon(\tau - 1)$. The base case of $\tau = 1$ clearly holds because $|I(\tau)| \leq 1$ and $C_\varepsilon(0) = 1$. Now consider epoch $\tau + 1$, assuming the claim holds for epoch τ , or more specifically $|I(\tau)| \leq 2\bar{L}_d C_\varepsilon(\tau - 1)$. If $x_0 \in [\underline{d}(\tau), \bar{d}(\tau)]$ then clearly $|I(\tau + 1)| \leq 2\bar{L}_d C_\varepsilon(\tau)$ by definition. For the other case of $x_0 \notin [\underline{d}(\tau), \bar{d}(\tau)]$ we have that $|I(\tau + 1)| = |I(\tau)|/2 \leq \bar{L}_d C_\varepsilon(\tau - 1)$. Note that $C_\varepsilon(\tau) \geq C_\varepsilon(\tau - 1)/2$. The lemma is thus proved. \square

EC.4.1. Proof of Theorem 2

We first consider the regret incurred by potential running out of inventory (at the end of the T total selling periods). Recall the definition that $\delta_t = \max\{0, f_0(p_t) - x_0\}$. By the Lipschitz continuity of f_0 , we have that

$$\sum_{t \leq T} \delta_t \leq \sum_{t \leq T} |f_0(p_t) - x_0| \leq \sum_{t \leq T} \bar{L}_p |p_t - p^c|.$$

Invoking Proposition 1, the regret of Algorithm 2 can be upper bounded by

$$\begin{aligned} \mathfrak{R}_{T, \varepsilon T}(\text{Alg. 2}; f_0, x_0) &\leq \mathbb{E} \left[\sum_{t=1}^T p^c x_0 - p_t f_0(p_t) + \bar{L}_p |p_t - p^c| \right] + \frac{1}{x_0} (\varepsilon T + \sqrt{T \ln T}) \\ &\leq (\max\{p^c \bar{L}_p, 1\} + \bar{L}_p) \times \mathbb{E} \left[\sum_{t=1}^T |p_t - p^c| \right] + \frac{1}{x_0} (\varepsilon T + \sqrt{T \ln T}). \end{aligned} \quad (\text{EC.25})$$

Here the second inequality holds because, if $p_t > p^c$ then $p^c x_0 - p_t f_0(p_t) \leq p^c (f_0(p^c) - f_0(p_t)) \leq p^c \times \bar{L}_p |p_t - p^c|$, and if $p_t < p^c$ then $p^c x_0 - p_t f_0(p_t) \leq (p^c - p_t) x_0 \leq |p^c - p_t|$.

We next upper bound the $\mathbb{E}[\sum_{t=1}^T |p_t - p^c|]$ term in Eq. (EC.25). We condition on the success event of $p^c \in I(\tau)$ for all epochs τ , which occurs with probability $1 - O(T^{-2} \log T)$ thanks to Lemma EC.5. Let τ_0 be the last complete epoch with positive inventory levels, satisfying that $2^{\tau_0} \leq T$. We then have that

$$\sum_{t=1}^T |p_t - p^c| \leq \sum_{\tau=1}^{\tau_0} T(\tau) \times |b(\tau) - a(\tau)| \leq \sum_{\tau=1}^{\tau_0} 2^\tau \times 2\bar{L}_d C_\varepsilon (\tau - 1) \quad (\text{EC.26})$$

$$\begin{aligned} &\leq \sum_{\tau=1}^{\tau_0} 2\bar{L}_d \left(\varepsilon T + 2^{\tau/2} \sqrt{\ln(2T^2)} + 2 \ln(2T^2) \right) \\ &\leq 2\bar{L}_d \varepsilon T \log_2 T + \frac{2\sqrt{2T \ln(2T^2)}}{\sqrt{2} - 1} + 2\bar{L}_d \ln(2T^2) \log_2 T. \end{aligned} \quad (\text{EC.27})$$

Here Eq. (EC.26) holds thanks to Lemma EC.5, and the last inequality holds because $2^{\tau_0} \leq T$.

Combining Eqs. (EC.25) and (EC.27) we have that

$$\begin{aligned} &sR_{T, \varepsilon T}(\text{Alg. 2}; f_0, x_0) \\ &\leq O(1) + \frac{1}{x_0} (\varepsilon T + \sqrt{T \ln T}) + (\max\{p^c \bar{L}_p, 1\} + \bar{L}_p) \left[2\bar{L}_d \varepsilon T \log_2 T + \frac{2\sqrt{2T \ln(2T^2)}}{\sqrt{2} - 1} + 2\bar{L}_d \ln(2T^2) \log_2 T \right] \\ &\leq (x_0^{-1} + 4\bar{p} \bar{L}_p \bar{L}_d) \varepsilon T \log_2 T + (x_0^{-1} + 14\bar{p} \bar{L}_p) \sqrt{T \ln(2T^2)} + 6\bar{p} \bar{L}_p \bar{L}_d \ln^2(2T^2) + O(1) \\ &= \tilde{O}(\varepsilon T + \sqrt{T}), \end{aligned}$$

which completes the proof of Theorem 2.

EC.5. Proofs of Technical Lemmas in Sec. 5.2

EC.5.1. Proof of Lemma 4

Fix a specific epoch τ and thread j such that $\varepsilon_j \geq \varepsilon$. Invoke Lemma EC.3 with $\delta = 1/T^2$, $T' = T(\tau)$ and $q = \wp_j$, we have that with probability $1 - O(T^{-2})$,

$$\begin{aligned} \left| \frac{\widehat{d}_j(\tau)}{\wp_j T(\tau)} - f_0(p_j(\tau)) \right| &\leq \min \left\{ 1, \frac{\varepsilon T}{T(\tau)} \right\} + \sqrt{\frac{\log(2T^2)}{\wp_j T(\tau)}} + \frac{2\log(2T^2)}{\wp_j T(\tau)} \\ &\leq \min \left\{ 1, \frac{\varepsilon_j T}{T(\tau)} \right\} + \sqrt{\frac{\log(2T^2)}{\wp_j T(\tau)}} + \frac{2\log(2T^2)}{\wp_j T(\tau)}, \end{aligned}$$

where the second inequality holds because $\varepsilon \leq \varepsilon_j$. The first property of Lemma 4 is then proved, by the definition of $C_{\varepsilon_j}(\tau)$ in Algorithm 3 and that $\wp_j T(\tau) = T_j(\tau)$.

The second and third properties can be proved in the same vein as the proof of Lemma EC.5, via an induction argument with the union bound over all epochs and threads.

EC.5.2. Proof of Corollary 1

According to Algorithm 3, J only decreases when $I_J(\tau) = \emptyset$. If $\varepsilon_J \leq \varepsilon$, then by Lemma 4 it holds that $p^c \in I_j(\tau)$ for all $j \leq J$ and τ . This means that J will never be further decreased since $I_J(\tau) \neq \emptyset$ throughout.

EC.6. Proofs of technical results in Sec. 6

EC.6.1. Proof of Lemma 5.

Invoking Lemma EC.3 with $T' = T_3$, $q = 1/N_3$, $\varepsilon T = ZT_3$ and $\delta = 1/(N_3 T^2)$, it holds with probability $1 - T^{-2}$ uniformly over all $i \in [N_3]$ that

$$\left| \frac{N_3 \widehat{d}(i)}{T_3} - f_0(p(i)) \right| \leq \min\{1, Z\} + \sqrt{\frac{N_3 \log(2N_3 T^2)}{T_3}} + \frac{2N_3 \log(2N_3 T^2)}{T_3} = H_Z(N_3, T_3). \quad (\text{EC.28})$$

We now prove the upper bound on $|\widehat{p}^c - p^c|$. Recall that p^c is the unique solution to $f_0(p^c) = x_0$, or equivalently $p^c = f_0^{-1}(x_0)$. Because of the Lipschitz continuity of f_0 (Assumption 2) and the fact that $p(i)$ and $p(i+1)$ are $(\bar{p} - \underline{p})/N_3$ distance apart, there exists $i^\# \in [N_3]$ such that $|f_0(p(i^\#)) - x_0| \leq \bar{L}_p(\bar{p} - \underline{p})/N_3$. This implies that, with probability $1 - O(T^{-2})$,

$$\begin{aligned} |f_0(p(\widehat{i}^c)) - x_0| &\leq |\widehat{d}(\widehat{i}^c) - f_0(p(\widehat{i}^c))| + |\widehat{d}(\widehat{i}^c) - x_0| \\ &\leq |\widehat{d}(\widehat{i}^c) - f_0(p(\widehat{i}^c))| + |\widehat{d}(i^\#) - x_0| \end{aligned} \quad (\text{EC.29})$$

$$\begin{aligned}
&\leq |\tilde{d}(\hat{i}^c) - f_0(p(\hat{i}^c))| + |\tilde{d}(i^\#) - f_0(p(i^\#))| + |f_0(p(i^\#)) - x_0| \\
&\leq 2H_Z(N_3, T_3) + \bar{L}_p(\bar{p} - \underline{p})/N_3.
\end{aligned} \tag{EC.30}$$

Here, Eq. (EC.29) holds because \hat{i}^c minimizes $|\tilde{d}(i) - x_0|$ by definition, and the last inequality holds with probability $1 - O(T^{-2})$ by applying Eq. (EC.28). Consequently,

$$\begin{aligned}
|p(\hat{i}^c) - p^c| &\leq |f_0^{-1}(x_0 \pm [\bar{L}_p(\bar{p} - \underline{p})/N_3 + 2H_Z(N_3, T_3)]) - f_0^{-1}(x_0)| \\
&\leq \bar{L}_d \times \left[\frac{\bar{L}_p(\bar{p} - \underline{p})}{N_3} + 2H_Z(N_3, T_3) \right],
\end{aligned}$$

where the last inequality holds thanks again to Assumption (A1). This proves the upper bound on $|\hat{p}^c - p^c|$.

Next we prove the upper bound on $|\hat{p}^o - p^o|$. Let $d^o = f_0(p^o)$ and $i^* \in [N_3]$ be the index such that $|p(i^*) - p^o| \leq (\bar{p} - \underline{p})/N_3$. By the Lipschitz continuity of f_0 (see Assumption A2) this means that $|d(p(i^*)) - d^o| \leq \bar{L}_p(\bar{p} - \underline{p})/N_3$. Recall also the definition that $r(p) := pf_0(p)$ and $r(d) = df_0^{-1}(d)$. By the strong smoothness of $r(d)$ (see Assumption A3), we have that

$$|r(d(p(i^*))) - r(d^o)| \leq \frac{M^2}{2} |d(p(i^*)) - d^o|^2 \leq \frac{M^2 \bar{L}_p^2 (\bar{p} - \underline{p})^2}{2N_3^2}. \tag{EC.31}$$

On the other hand, \hat{i}^o is selected such that $p(\hat{i}^o)\tilde{d}(\hat{i}^o) \geq p(i^*)\tilde{d}(i^*)$. Therefore,

$$\begin{aligned}
r(p(i^*)) - r(p(\hat{i}^o)) &= p(i^*)f_0(p(i^*)) - p(\hat{i}^o)f_0(p(\hat{i}^o)) \\
&\leq \bar{p}|\tilde{d}(\hat{i}^o) - f_0(p(\hat{i}^o))| + \bar{p}|\tilde{d}(i^*) - f_0(p(i^*))| + p(i^*)\tilde{d}(i^*) - p(\hat{i}^o)\tilde{d}(\hat{i}^o) \\
&\leq \bar{p}|\tilde{d}(\hat{i}^o) - f_0(p(\hat{i}^o))| + \bar{p}|\tilde{d}(i^*) - f_0(p(i^*))| \leq 2\bar{p}H_Z(N_3, T_3).
\end{aligned} \tag{EC.32}$$

Combining Eqs. (EC.31) and (EC.32), we have that

$$r(d^o) - r(f_0(p(\hat{i}^o))) \geq \frac{M^2 \bar{L}_p^2 (\bar{p} - \underline{p})^2}{2N_3^2} + 2\bar{p}H_Z(N_3, T_3), \tag{EC.33}$$

By strong concavity of $r(d)$ (see Assumption A3), Eq. (EC.33) implies that

$$|d^o - f_0(p(\hat{i}^o))| \leq \frac{2}{\sigma^2} \sqrt{\frac{M^2 \bar{L}_p^2 (\bar{p} - \underline{p})^2}{2N_3^2} + 2\bar{p}H_Z(N_3, T_3)} \leq \frac{2}{\sigma^2} \left[\frac{M\bar{L}_p(\bar{p} - \underline{p})}{\sqrt{2}N_3} + \sqrt{2\bar{p}H_Z(N_3, T_3)} \right].$$

Subsequently, using the Lipschitz continuity of f_0^{-1} (see Assumption A2) we have that

$$|\hat{p}^o - p^o| = |f_0^{-1}(f_0(p(\hat{i}^o))) - f_0^{-1}(d^o)| \leq \bar{L}_d |f_0(p(\hat{i}^o)) - d^o|$$

$$\leq \frac{2\bar{L}_d}{\sigma^2} \left[\frac{M\bar{L}_p(\bar{p}-\underline{p})}{\sqrt{2}N_3} + \sqrt{2\bar{p}H_Z(N_3, T_3)} \right],$$

which is to be demonstrated.

The next lemma is an immediate corollary of Lemma 5, showing that (with high probability) the estimation errors of p^c and p^o are smaller relative to the gap between p^c and p^o , when the parameters in Lemma 5 are tuned appropriately.

COROLLARY EC.1. *Suppose $p^c \neq p^o$, and for some positive constants $\beta' > \gamma' > 0$, that $T_3 = T^{\beta'}$ and $N_3 = T^{\gamma'}$. Suppose also that the Z parameter defined in Lemma 5 satisfies $Z \leq \frac{1}{3} \max\{\frac{|p^o-p^c|}{40\bar{L}_d}, \frac{1}{2\bar{p}}(\frac{\sigma^2|p^o-p^c|}{40\bar{L}_d})^2\}$. Then there exists a polynomial function $\varphi(\log T, \bar{L}_d, \bar{L}_p, \bar{p}, 1/|p^o-p^c|)$, whose degrees depend on β', γ' , such that if $T \geq \varphi(\log T, \bar{L}_d, \bar{L}_p, \bar{p}, 1/|p^o-p^c|)$, then with probability $1 - O(T^{-2})$ it holds that*

$$|\hat{p}^c - p^c| + |\hat{p}^o - p^o| \leq 0.2|p^o - p^c|.$$

EC.6.2. Proof of Corollary EC.1.

Following Lemma 5, it suffices to prove, separately, that all of $\frac{\bar{L}_d\bar{L}_p(\bar{p}-\underline{p})}{N_3}$, $2\bar{L}_dH_Z(N_3, T_3)$, $\frac{2\bar{L}_dM\bar{L}_p(\bar{p}-\underline{p})}{\sqrt{2}\sigma^2N_3}$ and $2\sigma^{-2}\bar{L}_d\sqrt{2\bar{p}H_Z(N_3, T_3)}$ terms are upper bounded by $0.05|p^o - p^c|$. to simplify notations, we shall also denote $\Delta p := |p^o - p^c|$ throughout the rest of this proof.

First we consider the $\frac{\bar{L}_d\bar{L}_p(\bar{p}-\underline{p})}{N_3} \leq 0.05\Delta p$ constraint. Re-arranging the terms we have that $N_3 \geq \frac{\bar{L}_d\bar{L}_p(\bar{p}-\underline{p})}{0.05\Delta p}$. Since $N_3 = T^{\gamma'}$, the condition can be reduced to

$$T \geq \left[\frac{\bar{L}_d\bar{L}_p(\bar{p}-\underline{p})}{0.05\Delta p} \right]^{1/\gamma'}. \quad (\text{EC.34})$$

Second we consider the $\frac{2\bar{L}_dM\bar{L}_p(\bar{p}-\underline{p})}{\sqrt{2}\sigma^2N_3} \leq 0.05\Delta p$ constraint. Again, with $N_3 = T^{\gamma'}$, the constraint is satisfied if

$$T \geq \left[\frac{2\bar{L}_dM\bar{L}_p(\bar{p}-\underline{p})}{0.1\sigma^2\Delta p} \right]^{1/\gamma'}. \quad (\text{EC.35})$$

We next consider the constraints $2\bar{L}_dH_Z(N_3, T_3) \leq 0.05\Delta p$ and $2\sigma^{-2}\bar{L}_d\sqrt{2\bar{p}H_Z(N_3, T_3)} \leq 0.05\Delta p$. We first simplify both constraints as conditions involving $H_Z(N_3, T_3)$ only, as

$$H_Z(N_3, T_3) \leq \max \left\{ \frac{\Delta p}{40\bar{L}_d}, \frac{1}{2\bar{p}} \left(\frac{\sigma^2\Delta p}{40\bar{L}_d} \right)^2 \right\} =: \bar{H}.$$

By definition of $H_Z(N_3, T_3)$, it suffices to prove that $Z \leq \bar{H}/3$, $\sqrt{\frac{N_3 \log(2N_3 T^2)}{T_3}} \leq \bar{H}/3$ and $\frac{2N_3 \log(2N_3 T^2)}{T_3} \leq \bar{H}_3/3$. Note that $Z \leq \bar{H}/3$ holds directly from the condition imposed on Z in Corollary EC.1. On the other hand, noting that $T_3 = T^{\beta'}$, $T_3/N_3 = T^{\beta' - \gamma'}$ and $\log(2N_3 T^2) \leq 3\log(2T)$, the conditions can be reduced to

$$T \geq (3\sqrt{3\log(2T)/\bar{H}})^{2/(\beta' - \gamma')}; \quad (\text{EC.36})$$

$$T \geq (18\log(2T)/\bar{H})^{1/(\beta' - \gamma')}. \quad (\text{EC.37})$$

Corollary EC.1 is subsequently proved, by noting that the right-hand sides of all Eqs. (EC.34) through (EC.37) are low-degree polynomials of $\bar{L}_d, \bar{L}_p, \bar{p}, 1/|p^o - p^c|, M, 1/\sigma$ and $\log T$, with degrees depending only on α', β' and γ' .

EC.6.3. Proof of Theorem 4.

The key step in this proof is to establish the first epoch ζ after which the estimates \hat{p}^c, \hat{p}^o are (with high probability) consistent. By consistency, we mean that $|\hat{p}^c - p^c| + |\hat{p}^o - p^o| \leq 0.2|p^c - p^o|$, the consequence of Corollary EC.1 established in the previous section. It is easy to verify that, with this inequality, $p^c < p^o$ implies $\hat{p}^c < \hat{p}^o$ and vice versa. Furthermore, $\underline{p}^o = (\hat{p}^c - \hat{p}^o)/2$ satisfies $p^c \leq \underline{p}^o \leq p^o$. With these conditions, the regret upper bounds proved in Theorems 1 and 3 can be directly applied.

To show $|\hat{p}^c - p^c| + |\hat{p}^o - p^o| \leq 0.2|p^c - p^o|$ with probability $1 - O(T^{-2})$, we only need to prove the conditions in Corollary EC.1 are satisfied. Consider an arbitrary epoch ζ in Algorithm 5. By definition, $\mathbb{E}[|\mathcal{G}(\zeta)|] = \sqrt{T(\zeta)}$ and $|\mathcal{G}(\zeta)|$ is the sum of $T(\zeta)$ i.i.d. binary random variables. By multiplicative Chernoff bound and the union bound, if $T(\zeta) \geq T_0 \geq 16 \ln T$ then it holds with probability $1 - O(T^{-1})$ uniformly over all ζ that $|\mathcal{G}(\zeta)| \geq \sqrt{T(\zeta)}/2$. Hence, $T_3 \geq \sqrt{T(\zeta)}/2$ and $N_3 \geq [T(\zeta)]^{1/4}/2$. Additionally, because there are at most εT selling periods being corrupted in epoch τ , and $\mathcal{G}(\zeta)$ are selected uniformly at random, we have with probability $1 - O(T^{-2})$ that

$$Z = \frac{1}{|\mathcal{G}(\zeta)|} \sum_{t \in \mathcal{G}(\zeta)} \iota_t = \mathbb{E}[Z] + O(\sqrt{\mathbb{E}[Z^2]} \log T + \log T/T_3) \leq O(\varepsilon T/T(\zeta)).$$

To satisfy the condition $Z \leq \frac{1}{3} \max\{\frac{|p^o - p^c|}{40\bar{L}_d}, \frac{1}{2\bar{p}}(\frac{\sigma^2 |p^o - p^c|}{40\bar{L}_d})^2\} =: B_Z$, it suffices for ζ to be large enough such that $T(\zeta) \geq \varepsilon T/B_Z$. Because B_Z is a polynomial of problem parameters and is independent of T , it suffices that $T(\zeta) = \Omega(\varepsilon T)$. To satisfy the conditions that $T_3 = T^{\beta'}$ and $N_3 = T^{\gamma'}$ for some $0 < \gamma' < \beta'$, we will simply set $\beta' = 0.2$ and $\gamma' = 0.1$. Subsequently, $T(\zeta)$ must satisfy $\sqrt{T(\zeta)} \geq T^{0.2}$, $[T(\zeta)]^{1/4} = T^{0.1}$, or more specifically $T(\zeta) \geq T^{0.4}$, to allow $T_3 = T^{\beta'}$ and $N_3 = T^{\gamma'}$ to hold. Since $T_0 = \lceil \sqrt{T} \rceil$, we conclude that all ζ satisfies $T_3 = T^{\beta'}$ and $N_3 = T^{\gamma'}$.

Now let ζ^\sharp be the first epoch such that $T(\zeta^\sharp) = \Omega(\varepsilon T)$, such that $Z \leq \frac{1}{3} \max\{\frac{|p^o - p^c|}{40L_d}, \frac{1}{2\bar{p}}(\frac{\sigma^2|p^o - p^c|}{40L_d})^2\}$ holds on and after epoch ζ^\sharp . Per the above discussion, this implies that epochs later than ζ^\sharp will have consistent \hat{p}^c, \hat{p}^o estimates, and therefore the regret of π^c or π^o selected by Algorithm 5 can be upper bounded by $\tilde{O}(\varepsilon T + \sqrt{T})$. Let also ζ_0 be the last epoch, which must satisfy $|\mathcal{T}(\zeta_0)| \leq T$. The regret of Algorithm 5 can then be upper bounded as

$$\mathfrak{R}_{T, \varepsilon T}(\text{Alg. 4}; f_0, x_0) \leq \sum_{\zeta \leq \zeta^\sharp} T(\zeta) + \sum_{\zeta > \zeta^\sharp} \tilde{O}(\varepsilon T + \sqrt{T}) \quad (\text{EC.38})$$

$$\begin{aligned} &\leq O(T_0 2^{\zeta^\sharp}) + \tilde{O}(\varepsilon T \zeta_0) + \tilde{O}(\sqrt{T} \zeta_0) \\ &\leq O(\varepsilon T) + \tilde{O}(\varepsilon T) + \tilde{O}(\sqrt{T}) = \tilde{O}(\varepsilon T + \sqrt{T}). \end{aligned} \quad (\text{EC.39})$$

Here, in Eq. (EC.38) we apply the upper regret bounds for π^o and π^c , and Eq. (EC.39) holds because $|\mathcal{T}(\zeta^\sharp)| = 2^{\zeta^\sharp} T_0 \asymp \varepsilon T$ and $\zeta_0 = O(\log T)$. This proves Theorem 4.

EC.7. Proof of Theorem 5

It suffices to prove that $\mathfrak{R}_{T, \varepsilon T}(\pi; f_0, x_0) \geq C' \max\{\varepsilon T, \sqrt{T}\}$ for some constant $C' > 0$. It is a standard result $\mathfrak{R}_{T, 0}(\pi; f_0, x_0) \geq \Omega(\sqrt{T})$ when there are no outlier customers (Wang et al. 2014). On the other hand, by setting the first εT customers as outliers who never make any purchases, it is clear that $\mathfrak{R}_{T, \varepsilon T}(\pi; f_0, x_0) \geq \Omega(\varepsilon T)$ because the regret is defined as $T r^* - \mathbb{E}^\pi[\sum_{t=1}^T r_t]$ where $r^* > 0$ is the expected per-period revenue of typical customers. This complete the proof of Theorem 5.