F.7 Explicit Bound

Assumptions. Before presenting the explicit bound, we introduce some assumptions to simplify the notation, though the results hold under more general conditions.

- **(A0)** We assume $\Sigma_a = \sigma_0^2 I_d$, $\Sigma = \tau^2 I_{d'}$, $\phi(x) = x$, $\|x\|_2 \leq 1$, and the matrices W_a are normalized such that $\lambda_1(\mathrm{W}_a \mathrm{W}_a^\top) = \lambda_d(\mathrm{W}_a \mathrm{W}_a^\top) = 1$.

Assumption **(A0)** can be relaxed; the theory holds for any positive definite matrices Σ_a and Σ, bounded L_2 norm contexts x, and any matrices W_a. Additionally, we make the following necessary assumptions:

- **(A1)** Let $G = \mathbb{E}_{X\sim\nu}[XX^\top]$ with $g = \lambda_d(G)$. We assume that $g > 0$.
- **(A2)** The fourth moment of X is bounded, and X satisfies $\sqrt{\mathbb{E}[(v^\top XX^\top v)^2]} \leq \mathrm{h} v^\top G v$ for all $v \in \mathbb{R}^d$. Additionally, X and $X \mid A$ follow the same distribution.

Theorem F.4 (Explicit Bound)**.** *Let $\pi_*(x)$ be the optimal action for context x. Under **(A0)**, **(A1)** and **(A2)**, the BSO of* `sDM` *under the structured prior (9) satisfies*

$$\mathrm{Bso}(\hat{\pi}_{\mathrm{G}}) \leq 2\sqrt{\mathbb{E}_{X\sim\nu}\left[\frac{d}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\tau^2\sigma_0^{-4} d}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2} + (\sigma_0^2 + \tau^2) d \exp\Big(-\frac{n\pi_0^2(\pi_*(X)|X)}{2}\Big)\right] + \frac{2d(\sigma_0^2+\tau^2)}{n}}\,,$$

where $\alpha_x = \lfloor \frac{\pi_0(\pi_*(x)|x)}{2} n \rfloor - 7\mathrm{h}\sqrt{\lfloor \frac{\pi_0(\pi_*(x)|x)}{2} n \rfloor (d + 2\ln(n))}$ for any $x \in \mathcal{X}$ and $a \in \mathcal{A}$.

Scaling with n. Note that

$$\alpha_X = \lfloor \frac{\pi_0(\pi_*(x)|x)}{2} n \rfloor - 7\mathrm{h}\sqrt{\lfloor \frac{\pi_0(\pi_*(x)|x)}{2} n \rfloor (d + 2\ln(n))}\,, \qquad \forall (x,a) \in \mathcal{X}\times\mathcal{A}\,. \tag{45}$$

In particular, with enough data n large enough, there exists a constant $c > 0$ such that

$$\alpha_x \geq c\pi_0(\pi_*(x)|x)n\,, \qquad \forall (x,a) \in \mathcal{X}\times\mathcal{A}\,.$$

Therefore, for n large enough, the bound could be expressed as

$$\begin{aligned}\mathrm{Bso}(\hat{\pi}_{\mathrm{G}}) &= \mathcal{O}\left(\sqrt{\mathbb{E}_{X\sim\nu}\left[\frac{d}{\pi_0(\pi_*(X)|X)n+1} + \frac{d}{\left(\pi_0(\pi_*(X)|X)n+1\right)^2} + d\exp\Big(-n\pi_0^2(\pi_*(X)|X)\Big)\right] + \frac{d}{n}}\right)\,,\\ &= \mathcal{O}\left(\sqrt{\mathbb{E}_{X\sim\nu}\left[\frac{d}{\pi_0(\pi_*(X)|X)n+1} + d\exp\Big(-n\pi_0^2(\pi_*(X)|X)\Big)\right] + \frac{d}{n}}\right)\,,\end{aligned}$$

where we omitted constants and used that $\frac{d}{(\pi_0(\pi_*(X)|X)n+1)^2} = \mathcal{O}(\frac{d}{\pi_0(\pi_*(X)|X)n+1})$ leads to the scaling provided in Theorem 5.2.

Proof idea. From Theorem 5.1, we establish that

$$\mathrm{Bso}(\hat{\pi}_{\mathrm{G}}) \leq 2\sqrt{d}\mathbb{E}\left[\|X\|_{\hat{\Sigma}_{\pi_*(X)}}\right]\,,$$

where we assume that $\phi(x) = x$. Thus we need to control the scaling of the posterior covariance matrix $\hat{\Sigma}_{\pi_*(X)}$ with n. First, note that for a fixed context x, the only randomness in $\hat{\Sigma}_{\pi_*(x)}$ originates from the design matrix $\hat{G}_{\pi_*(x)}$. Previous works simplify this by assuming the data is well-explored [Hong et al., 2023], i.e., $\hat{G}_a \succ \gamma n G$ for all $a \in \mathcal{A}$, where $G = \mathbb{E}_{X\sim\nu}\left[XX^\top\right]$. We instead only need to control $\hat{G}_{\pi_*(x)}$, and instead of assuming it satisfies $\hat{G}_{\pi_*(x)} \succ \gamma n G$, we will consider the events where this happens and the events where this is violated. We can do this based on when $\hat{G}_{\pi_*(x)}$ is the sum of n out-products. The problem, however, is that in this setting,

$\hat{G}_{\pi_*(x)}$ is the sum of N_a outer products, where the random variable N_a of the number of times the optimal action $\pi_*(X)$ appears in the sample set. To address this, we need to control N_a. Fortunately, we notice that $N_a|X$ follows a Binomial distribution with carefully chosen parameters. Using Hoeffding's inequality, we can bound the tails of this distribution, ensuring that the optimal action appears sufficiently often in the dataset. Once the optimal action counts are bounded, the results from Oliveira [2016] can be adapted to bound $\hat{G}_{\pi_*(x)}$. Matrix manipulations combined with eigenvalue inequalities then lead to the desired result.

Proof. From Theorem 5.1, we have that

$$\begin{aligned} \mathrm{BSO}(\hat{\pi}_{\mathrm{G}}) &\leq 2\sqrt{d}\mathbb{E}\left[\|\phi(X)\|_{\hat{\Sigma}_{\pi_*(X)}}\right], && (46)\\ &= 2\sqrt{d}\mathbb{E}\left[\|X\|_{\hat{\Sigma}_{\pi_*(X)}}\right], && (47)\\ &\leq 2\sqrt{d\,\mathbb{E}\left[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}}\right]}, && (48) \end{aligned}$$

where we used that the simplifying assumption that $\phi(x) = x$ in the second equality, and Cauchy-Schwartz in the third inequality. Now recall that to simplify, we also assumed that $\Sigma_a = \sigma_0^2 I_d$ for any $a \in \mathcal{A}$ and that $\Sigma = \tau^2 I_{d'}$. As a result, we have that:

$$\hat{\Sigma}_a = \tilde{\Sigma}_a + \sigma_0^{-4}\tilde{\Sigma}_a \mathrm{W}_a \bar{\Sigma} \mathrm{W}_a^\top \tilde{\Sigma}_a\,,$$

and we also have that $\lambda_1(\tilde{\Sigma}_a) \leq \sigma_0^2$ and that

$$\lambda_1(\hat{\Sigma}_a) \leq \sigma_0^2 + \tau^2\,, \qquad \forall a \in \mathcal{A}\,, \tag{49}$$

since $\lambda_1(\mathrm{W}_a\mathrm{W}_a^\top) = 1$. These are obtained using Weyl's inequalities.

Now we introduce some quantities:

- N_a : number of samples where $A_i = a$.
- $\epsilon_{x,a} = \mathbb{P}(A = a \mid x) = \pi_0(a \mid x)$ is the probability of choosing the action a given x.

Remark that $n - N_a \mid X \sim \mathrm{Bin}(n, 1 - \epsilon_{X,a})$. Therefore, via Hoeffding inequality, we can prove that for any $t > 0$ and $a \in \mathcal{A}$

$$\mathbb{P}(n - N_a - n(1 - \epsilon_{X,a}) > t \mid X) \leq \exp(-2t^2/n)\,,$$

which simplifies to

$$\mathbb{P}(N_a < n\epsilon_{X,a} - t \mid X) \leq \exp(-2t^2/n)\,.$$

In particular, for $t = n\epsilon_{X,a}/2$, we have that

$$\mathbb{P}(N_a < n\epsilon_{X,a}/2 \mid X) \leq \exp(-n\epsilon_{X,a}^2/2)\,.$$

If we let $\gamma_{x,a} = \epsilon_{x,a}/2$, we get that

$$\mathbb{P}(N_a < \gamma_{X,a} n \mid X) \leq \exp(-2n\gamma_{X,a}^2)\,.$$

Now, we define $\Omega_{x,a} = \{N_a \geq \gamma_{x,a} n\}$, then we have that

$$\mathbb{P}(\bar{\Omega}_{X,a} \mid X) \leq \exp(-2n\gamma_{X,a}^2)\,. \tag{50}$$

In particular,

$$\mathbb{P}(\bar{\Omega}_{X,\pi_*(X)} \mid X) \leq \exp(-2n\gamma_{X,\pi_*(X)}^2)\,. \tag{51}$$

First, we have that

$$
\begin{aligned}
\sqrt{\mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}}\big]} &= \sqrt{\mathbb{E}\big[\mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mid X\big]\big]}\,, \\
&= \sqrt{\mathbb{E}\big[\mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mathbb{I}\{\Omega_{X,\pi_*(X)}\} \mid X\big]\big] + \mathbb{E}\big[\mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mathbb{I}\{\bar{\Omega}_{X,\pi_*(X)}\} \mid X\big]\big]}\,, \\
&= \sqrt{\mathbb{E}\big[I_1\big] + \mathbb{E}\big[I_2\big]}\,.
\end{aligned} \tag{52}
$$

where $I_1 = \mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mathbb{I}\{\Omega_{X,\pi_*(X)}\} \mid X\big]$ and $I_2 = \mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mathbb{I}\{\bar{\Omega}_{X,\pi_*(X)}\} \mid X\big]\big]$.

To bound I_2, we use the assumption that $\|x\|_2 \leq 1$ for any context x and (49) leads to

$$
\|x\|^2_{\hat{\Sigma}_a} \leq \sigma_0^2 + \tau^2\,, \qquad\qquad \forall x, a\,.
$$

In particular, we have that

$$
\|x\|^2_{\hat{\Sigma}_{\pi_*(x)}} \leq \sigma_0^2 + \tau^2\,.
$$

Let $c_1 = \sigma_0^2 + \tau^2$. Then we have that

$$
\begin{aligned}
I_2 &= \mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}} \mathbb{I}\{\bar{\Omega}_{X,\pi_*(X)}\} \mid X\big]\,, \\
&\leq \mathbb{E}\big[c_1 \mathbb{I}\{\bar{\Omega}_{X,\pi_*(X)}\} \mid X\big]\,, \\
&= c_1 \mathbb{E}\big[\mathbb{I}\{\bar{\Omega}_{X,\pi_*(X)}\} \mid X\big]\,, \\
&= c_1 \mathbb{P}(\bar{\Omega}_{X,\pi_*(X)} \mid X)\,, \\
&\leq c_1 \exp(-2n\gamma^2_{X,\pi_*(X)})
\end{aligned} \tag{53}
$$

Lemma F.5 (Oliveira [2016]). *Assume $X_1, \ldots, X_n \in \mathbb{R}^d$ are i.i.d. random variables with bounded fourth moments. Define $G \equiv \mathbb{E}\left[X_1 X_1^\top\right]$ and $\widehat{G}_n \equiv \sum_{i=1}^n X_i X_i^\top$. Let $\mathrm{h} \in (1, +\infty)$ be such that $\sqrt{\mathbb{E}\left[\left(v^T X_1 X_1^\top v\right)^2\right]} \leq \mathrm{h} v^T G v$ for all $v \in \mathbb{R}^d$. Then for any $\delta \in (0,1)$:*

$$
\mathbb{P}\left(\widehat{G}_n \succeq \alpha(n,\delta) G\right) \geq 1 - \delta\,,
$$

where $\alpha(n,\delta) = n - 7\mathrm{h}\sqrt{n(d + 2\ln(2/\delta))}$. In particular, for any fixed positive definite matrix Λ, we have that

$$
\mathbb{P}\left((\widehat{G}_n + \Lambda)^{-1} \preceq (\alpha(n,\delta) G + \Lambda)^{-1}\right) \geq 1 - \delta\,, \tag{54}
$$

Let us focus on the firm term I_1. Thanks to multiplication with $\mathbb{I}\{\Omega_{X,\pi_*(X)}\}$, we know that given X, $N_{\pi_*(X)} \geq \gamma_{X,\pi_*(X)}\, n$. Therefore, given X, we have that

$$
\hat{G}_{\pi_*(X)} = \sigma^{-2} \sum_{i \in [n]} \mathbb{I}_{\{A_i = \pi_*(X)\}} X_i X_i^\top \succeq \sigma^{-2} \sum_{i=1}^{\lfloor \gamma_{X,\pi_*(X)}\, n \rfloor} X_i' X_i'^\top\,, \qquad\qquad \forall a \in \mathcal{A}\,, \tag{55}
$$

where the X_i' are copies of X_i (assuming that $X|A$ and X have the same law). To ease exposition, let $\alpha(\lfloor \gamma_{x,\pi_*(x)} n \rfloor, \delta) = \alpha_x$. Then, applying (54) in Lemma F.5 with $\Lambda = \Sigma_0^{-1} = \sigma_0^{-2} I_d$ and δ leads to

$$
\mathbb{P}\left((\hat{G}_{\pi_*(X)} + \sigma_0^{-2} I_d)^{-1} \mathbb{I}\{\Omega_{X,\pi_*(X)}\} \preceq (\sigma^{-2}\alpha_X G + \sigma_0^{-2} I_d)^{-1} \mid X\right) \geq 1 - \delta\,, \tag{56}
$$

where $G = \mathbb{E}_{X \sim \nu}\left[X X^\top\right]$. Noticing that $\tilde{\Sigma}_{\pi_*(X)} = (\hat{G}_{\pi_*(X)} + \sigma_0^{-2} I_d)^{-1}$ leads to

$$
\mathbb{P}\left(\tilde{\Sigma}_{\pi_*(X)} \mathbb{I}\{\Omega_{X,\pi_*(X)}\} \leq (\sigma^{-2}\alpha_X G + \sigma_0^{-2} I_d)^{-1} \mid X\right) \geq 1 - \delta\,. \tag{57}
$$

In particular, we bound its maximum eigenvalue as

$$\mathbb{P}\left(\lambda_1(\tilde{\Sigma}_{\pi_*(X)})\mathbb{I}\{\Omega_{X,\pi_*(X)}\} \leq \frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} \mid X\right) \geq 1 - \delta\,. \tag{58}$$

where $g = \lambda_d(G)$.

Moreover, we know that:

$$\hat{\Sigma}_a = \tilde{\Sigma}_a + \sigma_0^{-4}\tilde{\Sigma}_a \mathrm{W}_a \bar{\Sigma} \mathrm{W}_a^\top \tilde{\Sigma}_a\,,$$

and thus

$$\lambda_1(\hat{\Sigma}_a) = \lambda_1(\tilde{\Sigma}_a) + \sigma_0^{-4}\tau^2\lambda_1(\tilde{\Sigma}_a)^2\,,$$

where we use Weyl's inequality and the fact that for any matrix A and any positive semi-definite matrix B such that the product $\mathrm{A}^\top \mathrm{B}\mathrm{A}$ exists, the following inequality holds $\lambda_1(\mathrm{A}^\top\mathrm{B}\mathrm{A}) \leq \lambda_1(\mathrm{B})\lambda_1(\mathrm{A}^\top\mathrm{A})$. Therefore, from (58), we have that

$$\mathbb{P}\left(\lambda_1(\hat{\Sigma}_{\pi_*(X)})\mathbb{I}\{\Omega_{X,\pi_*(X)}\} \leq \frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\sigma_0^{-4}\tau^2}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2} \mid X\right) \geq 1 - \delta\,. \tag{59}$$

Using the law of total expectation, we get that

$$\begin{aligned} I_1 &= \mathbb{E}\big[\|X\|^2_{\hat{\Sigma}_{\pi_*(X)}}\mathbb{I}\{\Omega_{X,\pi_*(X)}\} \mid X\big]\,,\\ &\leq \frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\sigma_0^{-4}\tau^2}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2} + c_1\delta\,. \end{aligned}$$

Setting $\delta = 2/n$ leads to

$$I_1 \leq \frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\sigma_0^{-4}\tau^2}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2} + \frac{2c_1}{n}\,. \tag{60}$$

Combining (53) and (60) yields

$$\begin{aligned} \mathbb{E}\big[\|X\|_{\hat{\Sigma}_{\pi_*(X)}}\big] &\leq \sqrt{\mathbb{E}\big[I_1\big] + \mathbb{E}\big[I_2\big]}\,, \qquad (61)\\ &\leq \sqrt{\mathbb{E}_{X\sim\nu}\left[\frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\sigma_0^{-4}\tau^2}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2}\right] + \frac{2c_1}{n} + c_1\mathbb{E}_{X\sim\nu}\Big[\exp(-2n\gamma^2_{X,\pi_*(X)})\Big]}\,, \end{aligned}$$

Replacing $c_1 = \sigma_0^2 + \tau^2$ leads to

$$\begin{aligned} &\mathbb{E}\big[\|X\|_{\hat{\Sigma}_{\pi_*(X)}}\big] \qquad (62)\\ &\leq \sqrt{\mathbb{E}_{X\sim\nu}\left[\frac{1}{\sigma^{-2} g\alpha_X + \sigma_0^{-2}} + \frac{\sigma_0^{-4}\tau^2}{\left(\sigma^{-2} g\alpha_X + \sigma_0^{-2}\right)^2}\right] + \frac{2(\sigma_0^2+\tau^2)}{n} + (\sigma_0^2+\tau^2)\mathbb{E}_{X\sim\nu}\Big[\exp(-2n\gamma^2_{X,\pi_*(X)})\Big]}\,, \end{aligned}$$

Finally, noticing that $\gamma_{X,\pi_*(X)} = \frac{\pi_0(\pi_*(X)|X)}{2}$ concludes the proof. $\square$