We present the detailed proof of the result here. For some lemmas we follow the lines of Bastani and Bayati [2020] which prove an analogous bound for the Lasso estimator. For some calculation difference, we present them as well. We indicate it in the corresponding lemmas.

**Proof of Lemma 1.** Using $E[X'X \mathbf{1}_{(X \in U)}]$ is semi-positive definite,

$$E[X'X \mathbf{1}_{(X \in U)}] = E[X'X \mathbf{1}_{(X \in U)}] \cdot \frac{1}{P(x \in U)} \lessgtr E[X'X \mathbf{1}_{(X \in U)}] \cdot \frac{1}{p} \lessgtr E[X'X \mathbf{1}_{(X \in U)}] \cdot \frac{1}{p} + E[X'X \mathbf{1}_{(X \notin U)}] \cdot \frac{1}{p} = E[X'X] \cdot \frac{1}{p}.$$  

The following Lemma[A] states that the size of the set $T_{i,t}$ is $O(\log T)$.

**Lemma A (Lemma EC.8 of Bastani and Bayati [2020]).** When $t \geq (Kq)^2$, $Kq \geq 4$,

$$\frac{1}{2}q \log t < |T_{i,t}| < 2q \log t.$$

**Proof of Lemma[A]** We follow the lines of Lemma EC.8 of Bastani and Bayati [2020]. Let $N_t$ be the largest integer with $t > 2^{N_t+1}Kq$. Then $t \leq 2^{N_t+2}Kq$ and

$$(N_t + 2)q \leq |T_{i,t}| \leq (N_t + 3)q.$$  

For the lower bound, we have

$$\frac{\log(t/Kq)}{\log 2} < N_t + 2.$$  

Hence,

$$|T_{i,t}| \geq q \frac{\log(t/Kq)}{\log 2} \geq q \log(t/\sqrt{t}) = \frac{1}{2}q \log t.$$  

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).
The second inequality follows from $t > (Kq)^2$. For the upper bound, using $N_t + 1 \leq \frac{\log(t/Kq)}{\log 2}$,

$$|T_{i,t}| \leq \frac{(\log(t/Kq) + 2)q}{\log 2} = \frac{(\log(t/Kq) + \log 4)q}{\log 2} = \frac{(\log(4t/Kq))q}{\log 2} \leq 2q \log t.$$ 

The last inequality follows from $Kq \geq 4$.

Proof of Lemma \[4\] We follow the lines of Proposition 2 of Bastani and Bayati [2020]. By the Theorem \[2\] we have

$$\Pr \left( \lambda_{\min} \left( \hat{\Sigma}(T_{i,t}) \right) \leq \frac{\gamma p}{2} \right) \leq d \exp \left( -\frac{|T_{i,t}| \gamma p}{8} \right).$$

The size of the set $T_{i,t}$ is bounded by

$$|T_{i,t}| \geq \frac{1}{2} q \log t \geq \frac{8}{\gamma p} \log \left( \frac{t^2 d}{\alpha} \right),$$

provided that $q \geq \frac{8\gamma}{\gamma t}$ and $t \geq \frac{d}{\alpha}$. Hence, with probability at least $1 - \frac{\alpha}{t^2}$,

$$\lambda_{\min} \left( \hat{\Sigma}(T_{i,t}) \right) \geq \frac{\gamma p}{2}. \tag{1}$$

When $q \geq \frac{192d}{\gamma p} t^{1/2}$ and $t > \frac{2d+1}{\alpha}$, $|T_{i,t}| \geq 32\lambda_{\min}^{-1} \left( \hat{\Sigma}(T_{i,t}) \right) t^{1/2} \log(t^2(2d + 1)/\alpha)$. Then, Theorem \[1\] can be directly applicable with $\tau = \tau_0 \left( |T_{i,t}|/\log(t^2(2d + 1)/\alpha) \right)^{1/(1+\delta)}$, $\tau_0 \geq \nu_\delta$. Hence,

$$\Pr \left( \|\hat{\beta}(T_{i,t}) - \beta_i\|_2 \leq \left( \frac{\log(t^2(2d + 1)/\alpha)}{|T_{i,t}|} \right)^{\delta/(1+\delta)} \cdot 4\lambda_{\min}^{-1} \left( \hat{\Sigma}(T_{i,t}) \right) \tau_0 t^{1/2} \right) \geq 1 - \frac{\alpha}{t^2}.$$ 

Together with \[1\], when $q \geq 6 \left( \frac{32d t^{1/2}}{K\gamma p} \right)^{(1+\delta)/\delta}$ and $t \geq \frac{2d+1}{\alpha}$, with probability at least $1 - \frac{\alpha}{t^2}$,

$$\|\hat{\beta}(T_{i,t}) - \beta_i\|_2 \leq \frac{h}{4}.$$

Proof of Lemma \[7\] We follow the lines of Lemma EC.14 of Bastani and Bayati [2020]. We have

$$\mathbb{I}_{(r \in A_{i,t})} = \mathbb{I}_{(A_{r-1,t})} \cdot \mathbb{I}_{(x_r \in U_{i,t})} \cdot \mathbb{I}_{(r \notin \bigcup_{i \in [K]} T_{i,t})}.$$ 

For $n = 0, 1, 2, ...,$

$$r \in [(2^n - 1)Kq + 1, 2^n Kq]$$

are forced-sampling time steps and

$$r \in [2^n Kq + 1, (2^{n+1} - 1)Kq]$$

are not. Let $N_t$ be the largest integer such that $t > 2^{N_t+1} Kq$ as before. Define the intervals

$$V_{1,t} = [2^{N_t} Kq + 1, (2^{N_t+1} - 1)Kq], \quad V_{2,t} = [2^{N_t+1} Kq + 1, t \wedge (2^{N_t+2} - 1)Kq],$$

and the sum of random variables

$$M_{i,t} : = \sum_{r \in V_{1,t}} \mathbb{I}_{(r \in A_{i,t})} + \sum_{r \in V_{2,t}} \mathbb{I}_{(r \in A_{i,t})}$$

$$< \sum_{r = 1}^{N_t+1} \mathbb{I}_{(r \in A_{i,t})} = |A_{i,t}|.$$
Both intervals \( V_{1,t} \) and \( V_{2,t} \) are not containing the forced-sampling time steps and hence we do not update the forced-sample estimator within the intervals. Therefore, we can write

\[
M_{i,t} = \sum_{r \in V_{1,t}} \mathbb{I}(A_{2N_i Kq}) \cdot \mathbb{I}(x_r \in U_i) + \sum_{r \in V_{2,t}} \mathbb{I}(A_{2N_i + 1 Kq}) \cdot \mathbb{I}(x_r \in U_i)
\]

\[
\geq \mathbb{I}(A_{2N_i Kq}) \cdot \mathbb{I}(A_{2N_i + 1 Kq}) \cdot \sum_{r \in V_{1,t} \cup V_{2,t}} \mathbb{I}(x_r \in U_i).
\]

The lower bound of cardinality of two disjoint intervals is

\[
|V_{1,t} \cup V_{2,t}| = (t \wedge 2^{N_i+2} - 1) Kq - 2^{N_i+1} Kq + (2^{N_i+t+1} Kq - Kq - 2^{N_i} Kq)
\]

\[
= (t - 2^{N_i} Kq - Kq) \wedge (3 \cdot 2^{N_i} Kq - 2 Kq)
\]

\[
> \left( \frac{t}{2} - Kq \right) \wedge \left( \frac{3}{4} t - 2 Kq \right)
\]

\[
> \left( \frac{t}{2} - \frac{t}{80} \right) \wedge \left( \frac{3}{4} t - \frac{t}{40} \right)
\]

\[
= \frac{39}{80} t.
\]

The first inequality follows from \( t \leq 2^{N_i+2} Kq \). The last inequality follows from \( t > (Kq)^2 \) and \( q > 80 \). The upper bound of the cardinality of two disjoint intervals is

\[
|V_{1,t} \cup V_{2,t}| < t - 2^{N_i} Kq - Kq
\]

\[
< t - \frac{t}{4} - Kq
\]

\[
< \frac{3}{4} t.
\]

The probability of two events is bounded by

\[
P(A_{2N_i Kq} \text{ and } A_{2N_i+1 Kq}) \geq 1 - \frac{2K \alpha}{(t/4)^2} - \frac{2K \alpha}{(t/2)^2}
\]

\[
= 1 - \frac{32K \alpha}{t^2}
\]

\[
> 1 - 0.01.
\]

The last inequality is from \( t^2 > (Kq)^4 \) and \( \alpha \in (0, 1) \). Hence, we have

\[
\mathbb{E}[M_{i,t}] \geq P(A_{2N_i Kq} \text{ and } A_{2N_i+1 Kq}) P|V_{1,t} \cup V_{2,t}| \geq 0.48 t p.
\]

The Hoeffding’s inequality implies,

\[
P(\mathbb{E}[M_{i,t}] - M_{i,t} \geq \eta^2) \leq \exp \left( -\frac{2\eta^2}{|V_{1,t} \cup V_{2,t}|} \right)
\]

\[
\leq \exp \left( -\frac{8\eta^2}{3t} \right).
\]

Let \( \eta = 0.23 t p \). Then

\[
P(M_{i,t} < 0.48 t p - 0.23 t p) \leq \exp \left( -\frac{8}{3} t (0.23 p)^2 \right)
\]

\[
\leq \exp(-t p^2/9).
\]

Since \( M_{i,t} \leq |A_{i,t}| \),

\[
P\left(|A_{i,t}| < \frac{tp}{4}\right) \leq \exp(-t p^2/9) \leq \frac{\alpha}{t^2},
\]

provided that \( t \geq \frac{1}{\alpha} \) and \( q \geq \frac{54}{p} \). \qed
We now provide the proof of the expected regret bound.

**Proof of Theorem** \[ \square \] Lemma EC.19 of [Bastani and Bayati 2020] states that the upper bound of expected regret can be decomposed into

\[
\sum_{t=1}^{T} \mathbb{E}[r_t] = \sum_{t} \mathbb{E}[x^T \beta_{a^*(t)} - x^T \beta_{a(t)}] \\
\leq 2 \sum_{i \in D} \mathbb{P}(\|\hat{\beta}(S_{a^*(t),t-1}) - \beta_{a^*(t)}\|_2 > \Delta) + 2 \sum_{i \in D} \mathbb{P}(\|\hat{\beta}(S_{a(t),t-1}) - \beta_{a(t)}\|_2 > \Delta) + 4\Delta^2 K C_0
\]

for \( \Delta > 0 \). From Lemma 7 with \( \alpha = (2d + 1)t \), we have

\[
\mathbb{P}\left(\|\hat{\beta}(S_{i,t}) - \beta_i\|_2 \geq \left(\frac{4}{pt \log t}\right)^{\delta/(1+\delta)} \frac{32\tau_0 d^{1/2}}{\gamma p}\right) \leq \frac{3(2d + 2)}{t}
\]

for \( i \in K_{opt} \). Let \( \Delta = \left(\frac{4}{pt \log t}\right)^{\delta/(1+\delta)} \frac{32\tau_0 d^{1/2}}{\gamma p} \) then,

\[
\mathbb{E}[r_t] \leq \frac{12K(2d + 1)}{t} + 4 \left(\frac{32\tau_0}{\gamma p}\right)^2 d \left(\frac{4}{pt \log T}\right)^{2\delta/(1+\delta)} K C_0.
\]

The cumulative regret is bounded by

\[
\sum_{t=1}^{T} \mathbb{E}[r_t] \leq 12K(2d + 1)(\log T + 1) + 4^7 d \left(\frac{\tau_0}{\gamma}\right)^2 \frac{1}{p^3} K C_0 ((\log T)^2 + \log T)
\]

when \( \delta = 1 \) and

\[
\sum_{t=1}^{T} \mathbb{E}[r_t] \leq 12K(2d + 1)(\log T + 1) + 64^2 16 \frac{\delta}{1+\delta} d \left(\frac{\tau_0}{\gamma}\right)^2 \frac{1}{p^{2+\delta}} K C_0 \left(\frac{1 + \delta}{1-\delta}\right) T \frac{\delta}{2+\delta} (\log T)^{\frac{2\delta}{1+\delta}}
\]

when \( 0 < \delta < 1 \).

\[ \square \]

**References**