Proof of Lemma 1

Recall the definition of oracle solution, Equation (2), and we have:

$$\hat{\boldsymbol{\beta}}_O = \arg \min_{\substack{\boldsymbol{\beta}_{S^c}=0 \\ \boldsymbol{\beta}_S \in S}} \|\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{y}\|_2^2 = \arg\min \|\boldsymbol{X}_S \boldsymbol{\beta}_S - \boldsymbol{y}\|_2^2, \tag{6}$$

where $\boldsymbol{X}_S$ denotes the sub-matrix of $\boldsymbol{X}$ that contains columns indexed by the non-zero index set S. The first order optimal condition for Equation (6) is:

$$\frac{1}{n}\boldsymbol{X}_S^T(\boldsymbol{X}_S\hat{\boldsymbol{\beta}}_O - \boldsymbol{y}) = 0$$

$$\Rightarrow \frac{1}{n}X_S^T(X_S\hat{\boldsymbol{\beta}}_0 - X_S\boldsymbol{\beta}_S^{true} + \boldsymbol{\epsilon}) = 0 \tag{7}$$

$$\Rightarrow \frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true}) = \frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon} \tag{8}$$

where (7) comes from the facts that $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}^{true} + \boldsymbol{\epsilon}$ and $\beta_{S^c}^{true} = 0$. We then multiply (8) by $(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})^T$.

$$\frac{1}{n}(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})^T\boldsymbol{X}_S^T\boldsymbol{X}_S(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true}) = \frac{1}{n}(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})^T\boldsymbol{X}_S^T\boldsymbol{\epsilon},$$

which leads to the following inequality:

$$\text{eig}_{\min}\left(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S\right)\|(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})\|_2^2 \le \frac{1}{n}(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})^T\boldsymbol{X}_S^T\boldsymbol{\epsilon} \le \|(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})\|_2\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|_2,$$

where the first inequality comes from the observation that $\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\|(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})\|_2^2 \le \frac{1}{n}(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})^T\boldsymbol{X}_S^T\boldsymbol{X}_S(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})$. Divided both sides by $\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\|(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})\|_2$, we will have

$$\|(\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true})\|_2 \le \frac{\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|_2}{\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}.$$

Since $\boldsymbol{\epsilon}$ is σ-subgaussian and $\boldsymbol{x}$ is upper bounded, to bound the $\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|$ term, we can use the Hanson-Wright inequality Rudelson et al. [2013]:

$$\mathbb{P}\{|\boldsymbol{v}^TA\boldsymbol{v} - \mathbb{E}[\boldsymbol{v}^TA\boldsymbol{v}]| > t\} \le 2\exp\left(-C_h\min\left\{\frac{t^2}{\sigma^4\|A\|_F^2}, \frac{t}{\sigma^2\|A\|_2}\right\}\right),$$

where $\boldsymbol{v} = (v_1, ..., v_n)$ and v_i is a zero mean σ-sub-Gaussian random variable. As $\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|_2 = \sqrt{\frac{1}{n}\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}}$ and $\boldsymbol{\epsilon}$ follows σ-sub-Gaussian distribution centered at zero, we can apply the Hanson-Wright inequality to $\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}$ by setting $t = \mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]$:

$$\begin{aligned}
&\mathbb{P}\left\{|\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon} - \mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]| > \mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]\right\} \\
\le& 2\exp\left(-C_h\min\left\{\frac{\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}}{\sigma^2\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S\|_2}, \frac{(\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon})^2}{\sigma^4\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S\|_F^2}\right\}\right) \\
\le& 2\exp\left(-C_h\min\left\{\frac{\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\text{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}\frac{E[\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}]}{\sigma^2}, \frac{\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}{\text{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}\frac{E[\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}]^2}{s\sigma^4}\right\}\right) \\
\le& 2\exp\left(-C_h\min\left\{\frac{\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\text{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}n, \frac{\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}{\text{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}\frac{n^2}{s}\right\}\right) \\
\le& 2\exp\left(-n\frac{C_h\text{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\text{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}\right).
\end{aligned}$$

The last inequality holds when $n \geq s\frac{\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\mathrm{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}$. Define the projection matrix $P_j = \boldsymbol{X}_S(\boldsymbol{X}_S^T\boldsymbol{X}_S)^{-1}\boldsymbol{X}_S^T$, then we have $(P_j\boldsymbol{\epsilon})^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)(P_j\boldsymbol{\epsilon}) = \boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}$. Conditioning on the event $\mathcal{E}_1 = \left\{|\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon} - \mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]| \leq \mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]\right\}$, the following inequality will be held:

$$\begin{aligned}
\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|_2 = \sqrt{\frac{1}{n}\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}} &\leq \sqrt{\frac{2}{n}\mathbb{E}[\boldsymbol{\epsilon}^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\boldsymbol{\epsilon}]} \\
&= \sqrt{\frac{2}{n}\mathbb{E}[(P_j\boldsymbol{\epsilon})^T(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)(P_j\boldsymbol{\epsilon})]} \\
&\leq \sqrt{\frac{2}{n}\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\mathbb{E}[\|P_j\boldsymbol{\epsilon}\|_2^2]} \\
&= \sqrt{2\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)\sigma^2\frac{s}{n}}
\end{aligned}$$

It is worths pointing out that P_j is the projection matrix that projects the n-dimensional vector $\boldsymbol{\epsilon}$ onto s-dimensional subspace and $\mathbb{E}[P_j\boldsymbol{\epsilon}] = 0$, which directly leads to $\mathbb{E}[\|P_j\boldsymbol{\epsilon}\|_2^2] = Var(P_j\boldsymbol{\epsilon}) = s\sigma^2$. We define

$$\delta_1 = 2\exp\left(-n\frac{C_h\mathrm{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}\right). \tag{9}$$

Therefore, with probability $1-\delta_1$, the following inequality holds:

$$\|\hat{\boldsymbol{\beta}}_O - \boldsymbol{\beta}^{true}\|_2 \leq \frac{\|\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{\epsilon}\|_2}{\mathrm{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)} \leq \sqrt{2\sigma^2\frac{\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\mathrm{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}\frac{s}{n}},$$

where $\delta_1 = \exp(-O(n))$. □

Proof of Proposition 1

We first expand and restate Proposition 1 as follows: If the compatibility condition is satisfied, the the error follows i.i.d σ-subgaussian distribution, and $\min\{|\beta_j^{true}| : \beta_j^{true} \neq 0,\ j = 1, 2, ..., d\} > (4\frac{s}{\phi^2} + a)\lambda$, then the following MCP inequality holds:

$$\mathbb{P}\left\{\|\hat{\boldsymbol{\beta}}_{2sWL}(\boldsymbol{X}, \boldsymbol{y}, \lambda) - \boldsymbol{\beta}^{true}\|_2 > \sqrt{\frac{2\sigma^2\mathrm{eig}_{\max}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)}{\mathrm{eig}_{\min}(\frac{1}{n}\boldsymbol{X}_S^T\boldsymbol{X}_S)^2}}\sqrt{\frac{s}{n}}\right\} \leq \delta_1 + \delta_2 + \delta_3$$

where δ_1 is defined before and

$$\begin{aligned}
\delta_2 &= 2d\exp\left(-\frac{n\lambda^2}{32x_{\max}^2\sigma^2}\right), \\
\delta_3 &= 2(d-s)\exp\left(-\frac{a^2n\lambda^2}{2x_{\max}^2\sigma^2}\right).
\end{aligned}$$

Note that Theorem 1 in Fan et al. [2014] states that if events $\mathcal{E}_2 = \{\|\boldsymbol{\beta}_1 - \boldsymbol{\beta}^{true}\|_\infty \leq 4\frac{s}{\phi^2}\lambda\}$ and $\mathcal{E}_3 = \{\|\frac{1}{n}X_{S^c}^T(\boldsymbol{X}_S\hat{\beta}_O - y)\|_\infty \leq a\lambda\}$ hold, then $\hat{\boldsymbol{\beta}}_{2sWL} = \hat{\boldsymbol{\beta}}_o$. Therefore, to prove this expanded version of Proposition 1, we merely need to derive probability bounds for those two events.

First, let us consider the $\mathcal{E}_2$. As in 2sWL algorithm, we solve for the lasso solution in the first step, i.e., $\boldsymbol{\beta}_1$ is a lasso solution. Under compatibility condition, Van De Geer et al. [2009] establishes that the following inequality holds with probability $1-\delta_2$:

$$\|\boldsymbol{\beta} - \boldsymbol{\beta}^{true}\|_1 \leq 4\frac{s}{\phi^2}\lambda$$

Combining with the fact that $\|\cdot\|_\infty \leq \|\cdot\|_1$, we have

$$\mathbb{P}\{\mathcal{E}_2\} \geq 1 - 2d\exp(-\frac{n\lambda^2}{32x_{\max}^2\sigma^2}) = 1 - \delta_2.$$

Now we consider event $\mathcal{E}_3$, whose probability can be bounded as follows:

$$\begin{aligned}
&\mathbb{P}\{\|\frac{1}{n}\boldsymbol{X}_{S^c}^T(\boldsymbol{X}_S\hat{\boldsymbol{\beta}}_O-\boldsymbol{y})\|_\infty \geq a\lambda\}\\
\leq &\sum_{i\in S^c}\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(\boldsymbol{X}_S\hat{\boldsymbol{\beta}}_O-y)\|_\infty \geq a\lambda\}\\
= &\sum_{i\in S^c}\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(\boldsymbol{X}_S(\boldsymbol{X}_S^T\boldsymbol{X}_S)^{-1}\boldsymbol{X}_S^T\boldsymbol{y}-\boldsymbol{y})\|_\infty \geq a\lambda\} \qquad (10)\\
= &\sum_{i\in S^c}\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(P_j\boldsymbol{y}-\boldsymbol{y})\|_\infty \geq a\lambda\}. \qquad (11)
\end{aligned}$$

where the first inequality comes from the union bound on non-significant dimensions, $\boldsymbol{x}_i$ is the ith column of covariate matrix $\boldsymbol{X}$, Equation (10) comes from the fact that Equation (6) allows a closed-form solution $\hat{\boldsymbol{\beta}}_O = \boldsymbol{X}_S(\boldsymbol{X}_S^T\boldsymbol{X}_S)^{-1}\boldsymbol{X}_S^T\boldsymbol{y}$, and Equation (11) directly follows $P_j = \boldsymbol{X}_S(\boldsymbol{X}_S^T\boldsymbol{X}_S)^{-1}\boldsymbol{X}_S^T$. As $P_j\boldsymbol{y} = \boldsymbol{X}_S(\boldsymbol{X}_S^T\boldsymbol{X}_S)^{-1}\boldsymbol{X}_S^T(\boldsymbol{X}_S\boldsymbol{\beta}_S^{true}+\boldsymbol{\epsilon}) = \boldsymbol{X}_S\boldsymbol{\beta}_S^{true}+P_j\boldsymbol{\epsilon}$, we can further simplify Equation (11) as follows:

$$\begin{aligned}
&\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(P_j\boldsymbol{y}-\boldsymbol{y})\|_\infty \geq a\lambda\}\\
=&\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(\boldsymbol{X}_S\boldsymbol{\beta}_S^{true}+P_j\boldsymbol{\epsilon}-\boldsymbol{X}_S\boldsymbol{\beta}_S^{ture}-\boldsymbol{\epsilon})\|_\infty \geq a\lambda\}\\
=&\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(P_j-I)\boldsymbol{\epsilon}\|_\infty \geq a\lambda\}.
\end{aligned}$$

We then apply Hoeffding's inequality and have the following results:

$$\begin{aligned}
\mathbb{P}\{\|\frac{1}{n}\boldsymbol{x}_i^T(P_j-I)\boldsymbol{\epsilon}\|_\infty \geq a\lambda\} \leq& 2\exp\left(-\frac{a^2\lambda^2}{2\sigma^2\|\frac{1}{n}\boldsymbol{x}_i^T(P_j-I)\|_2^2}\right)\\
=&2\exp\left(-\frac{n^2a^2\lambda^2}{2\sigma^2\boldsymbol{x}_i^T(I-P_j)(I-P_j)^T\boldsymbol{x}_i}\right)\\
\leq&2\exp\left(-\frac{na^2\lambda^2}{2\sigma^2x_{\max}^2}\right),
\end{aligned}$$

where the second inequality comes from the facts that $\boldsymbol{x}_i^2 \leq nx_{\max}^2$, where $x_{\max}\doteq\|\boldsymbol{x}\|_\infty$, and $(I-P_j)\preceq I$. Therefore, with union bound over $i \in S^c$, we can bound event $\mathcal{E}_3$:

$$\mathbb{P}\{\mathcal{E}_3\} \geq 1-2(d-s)\exp\left(-\frac{na^2\lambda^2}{2\sigma^2x_{\max}^2}\right) = 1-\delta_3.$$

Thus, with probability $(1-\delta_2-\delta_3)$, $\boldsymbol{\beta}_2$ derived from the 2sWL procedure is an oracle solution. Finally, combining with event $\mathcal{E}_1$ in Lemma 1, Proposition 1 follows immediately. □

Proof of Proposition 2

When $\lambda = O(\sqrt{\log d/n})$, the minimum sample size needed to satisfy Proposition 1 can be derived as follows. Let $\beta_{\min} = \min\{|\boldsymbol{\beta}_j^{true}| : \beta_j^{true} \neq 0, j = 1, 2, ..., d\}$ and we have:

$$\beta_{\min} > (4\frac{s}{\phi^2}+a)\lambda. \qquad (12)$$

As we require $\lambda = O(\sqrt{\log d/n})$, Equation (12) implies:

$$\begin{aligned}
&\beta_{\min} \gtrsim O\left((4\frac{s}{\phi^2}+a)\sqrt{\frac{\log d}{n}}\right)\\
\Rightarrow& n \gtrsim O\left(\frac{(4s+a\phi^2)^2}{\phi^4}\frac{\log d}{\beta_{\min}^2}\right) \gtrsim O(s^2\log d)
\end{aligned}$$

Therefore, if we have a sample size larger than $O(s^2 \log d)$, we have $\|\hat{\boldsymbol{\beta}}_{2sWL}(\boldsymbol{X}, \boldsymbol{y}, \lambda) - \boldsymbol{\beta}^{true}\|_2 \leq O(\sqrt{\frac{s}{n}})$ with certain probability (see Proposition 1), which implies that the cardinalities of $\hat{\boldsymbol{\beta}}_{2sWL}(\boldsymbol{X}, \boldsymbol{y}, \lambda)$ and $\boldsymbol{\beta}^{true}$ are s. Accordingly, we have $\|\hat{\boldsymbol{\beta}}_{2sWL}(\boldsymbol{X}, \boldsymbol{y}, \lambda) - \boldsymbol{\beta}^{true}\|_1 \leq O(\sqrt{s}\sqrt{\frac{s}{n}}) = O(s\sqrt{1/n})$. Comparing to the Lasso bound in Van De Geer et al. [2009], $\|\boldsymbol{\beta}_{Lasso} - \boldsymbol{\beta}^{true}\|_1 \leq O(s\sqrt{\frac{\log d}{n}})$, the convergence bound of the MCP estimator under the 2sWL procedure is faster, when the sample size is large enough (e.g. $n > O(s^2 \log d)$).

In addition, with high probability, the MCP estimator under the 2sWL procedure will match the oracle estimator, which mimics the scenario under which we solve the unpenalized problem with only significant dimensions. Therefore, the MCP estimator under the 2sWL procedure will match the oracle convergence rate with high probability. □

Proof of Theorem 1

There are four key steps in establishing the expected cumulative regret upper-bound for the MCP-Bandit algorithm in Theorem 1. The proofs for the first three steps (i.e., oracle inequality for non-i.i.d. data, oracle inequality for forced-sample estimator, and oracle inequality for all-sample estimator) have been detailed in §5.1, §5.2, and §5.3 in the main paper.

In the following, we will expand and provide details for the fourth step (i.e., bounding the cumulative expected regret) to complete the proof. We first divide our time periods $[T]$ into three groups:

1. $t \leq (Kq)^2$ with all samples and $t > (Kq)^2$ with forced samples;
2. $t > (Kq)^2$ without forced samples and the event $A_{t-1} \doteq \left\{\|\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{i,t-1}, \lambda) - \boldsymbol{\beta}^{true}\|_1 \leq \frac{h}{4x_{\max}}\right\}$ doesn't hold;
3. $t > (Kq)^2$ without forced samples and the event A_t holds.

The first group contains the forced samples and all samples with $t \leq (Kq)^2$. When $t \leq (Kq)^2$, we do not have sufficient samples to accurately estimate covariates parameter vectors, the decision performance under the MCP-Bandit algorithm will be sub-optimal comparing to that of the oracle case. Note that as we assume $\|x\|_\infty \leq x_{\max}$, $\|\boldsymbol{\beta}\|_1 \leq b$, each user's regret is bounded by $2bx_{\max}$ for any decision; and then the regret for $t \leq (Kq)^2$ is upper bounded by $(Kq)^2 2bx_{\max}$. Since the sampling frequency of forced sampling decays exponentially, there exists a constant $C_1 > 0$ such that the forced sample size $|\mathcal{T}_{i,t}|$ by time $t > (Kq)^2$ is bounded as follows:

$$\frac{1}{C_1} q \log t \leq |\mathcal{T}_{t,k}| \leq C_1 q \log t. \tag{13}$$

We, therefore, can bound the cumulative regret by $2(Kq)^2 bx_{\max} + 2C_1 qbx_{\max} \log T$.

The second group includes scenarios where $t > (Kq)^2$ and forced-sample-based estimators are not accurate enough. In particular, when A_{t-1} doesn't hold, the forced sample based estimator vector $\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{i,t}, \lambda)$ is not near the true parameter vector $\boldsymbol{\beta}^{true}$. Under those scenarios, our decisions will be sub-optimal with high probability. To bound the second group, we need to bound the expected instances that A_{t-1} doesn't hold from $(Kq)^2 < t \leq T-1$. According to Equation (13), for arm k, the forced sample size $|\mathcal{T}_{t-1,k}|$ is lower bounded by $q/C_1 \log t$ for time $t > (Kq)^2$. Since we require $q \gtrsim O(s^2 \log d) \geq \frac{8C_1 s\lambda_{\max}\sigma^2 x_{\max}}{h^2\lambda_{\min}}$, we can show that the following inequality holds for $T > 3$:

$$\begin{aligned} &\frac{\lambda_{\max}\sigma^2}{\lambda_{\min}} \cdot \frac{8sx_{\max}^2}{h^2} \leq \frac{1}{C_1} q \log T \leq |\mathcal{T}_{t,k}| \\ \Rightarrow &\sqrt{\frac{\lambda_{\max}\sigma^2}{2\lambda_{\min}}} \sqrt{\frac{s}{|\mathcal{A}|}} \leq \frac{h}{4x_{\max}}. \end{aligned}$$

Through union bounds over different arms and Proposition 3, there exist some positive constants C_4 and C_5 satisfying:

$$\sum_{t=(Kq)^2}^{T-1} \mathbb{P}\{A_t^c\} \leq \sum_{t=(Kq)^2}^{T-1} \sum_{k \in \mathcal{K}} \exp\left(-C_4 |\mathcal{T}_{t,k}| + C_5 \log d\right)$$
$$\leq \sum_{t=(Kq)^2}^{T-1} K \exp\left(-\frac{C_4}{C_1} q \log t + C_5 \log d\right).$$

when $q \geq \frac{C_1}{C_4}(C_5 \log d + 1)$, we have:

$$\sum_{t=(Kq)^2}^{T-1} \mathbb{P}\{A_t^c\} \leq K \sum_{t=(Kq)^2}^{T-1} \exp\left(-\log t\right) = K \sum_{t=(Kq)^2}^{T-1} \frac{1}{t} \leq K \log T, \tag{14}$$

where the last inequality uses the fact that $\sum_{t=(Kq)^2}^{T-1} \frac{1}{t} \leq \sum_4^T \frac{1}{t} \leq \log T$. Therefore, the expected regret bound for second group is $2Kbx_{\max} \log T$.

To bound the third group, without loss of generality, we assume that decision j is optimal (i.e., $j = \arg\max_{i \in [K]} X^T \boldsymbol{\beta}_i$). Then, the expected regret at time t is

$$r_t = \mathbb{E}\left(\sum_i \mathbf{1}[\text{choose decision } i] \cdot [\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i)]\right)$$
$$\leq \mathbb{E}\left(\sum_i \mathbf{1}[\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j] \cdot [\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i)]\right),$$

where the last inequality uses the fact that event $\{j = \arg\max_{i \in [K]} \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i\}$ is a subset of the event $\{\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j\}$ and that $\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i) \geq 0$. Thus, we can bound r_t through the regret incurred by each arm in K with respect to the optimal arm. We define the event $B_i = \{\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i) \geq 2\sqrt{s}\delta x_{\max}\}$, where $\delta = \sqrt{\frac{\sigma^2 \lambda_{\max} s}{2\lambda_{\min}^2 q^* t}}$. Then we can write

$$r_t \leq \mathbb{E}\left(\sum \mathbf{1}[(\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j) \cap B_i] \cdot [\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i)]\right)$$
$$+ \mathbb{E}\left(\sum \mathbf{1}[(\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j) \cap B_i^c] \cdot [\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i)]\right).$$

By the fact that $\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i) \leq 2bx_{\max}$ and the definition of B_i, we can further bound the regret as follows:

$$r_t \leq \sum 2bx_{\max} \mathbb{P}[(\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j) | B_i] + \sum 2\sqrt{s}\delta x_{\max} \mathbb{P}[B_i^c].$$

Note that if we choose decision i instead of decision j under the condition that event B_i happens, then the following inequality must hold:

$$0 > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j - \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i \geq \boldsymbol{x}_t^T(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + \boldsymbol{x}_t^T(\hat{\boldsymbol{\beta}}_i - \boldsymbol{\beta}_i) + 2\sqrt{s}\delta x_{\max}$$

Thus, at least, either $X_{t+1}^T(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) < -\sqrt{s}\delta x_{\max}$ or $X_{t+1}^T(\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i) \leq -\sqrt{s}\delta x_{\max}$. Therefore,

$$\begin{aligned}
&\mathbb{P}[(\boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_i > \boldsymbol{x}_t^T \hat{\boldsymbol{\beta}}_j) | B_i] \\
\leq &\mathbb{P}[\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}_j\|_1 \geq \sqrt{s}\delta] + \mathbb{P}[\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\|_1 \geq \sqrt{s}\delta] \\
\leq &\mathbb{P}[\sqrt{s}\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}\|_2 \geq \sqrt{s}\delta] + \mathbb{P}[\sqrt{s}\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\|_2 \geq \sqrt{s}\delta] \\
\leq &\mathbb{P}[\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}\|_2 \geq \delta] + \mathbb{P}[\|\boldsymbol{\beta}_i - \hat{\boldsymbol{\beta}}_i\|_2 \geq \delta].
\end{aligned} \tag{15}$$

Consider that we choose $\delta = \sqrt{\frac{\sigma^2 \lambda_{\max} s}{2\lambda_{\min}^2 q^* t}}$, and that we can find C_7, C_8, and $C_2 > 0$ via Proposition 5 such that

$$\mathbb{P}[\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}\|_2 \leq \delta\} \leq \exp\left(-C_7 t + C_8 \log d\right) + \frac{C_2}{t}, \quad j \in \mathcal{K}. \tag{16}$$

If we require $q \geq \frac{C_8+\sqrt{C_8^2+4C_7C_8\log d}}{2C_7}$ and $t \geq (Kq)^2$, then Equation (16) implies that

$$\mathbb{P}[\|\boldsymbol{\beta}_j - \hat{\boldsymbol{\beta}}\|_2 \geq \delta] \leq \exp\left(-\log t - C_8 \log d + C_8 \log d\right) + \frac{C_2}{t} = \frac{C_2+1}{t}. \tag{17}$$

Therefore, when $t \geq \left(\frac{C_8+\sqrt{C_8^2+4C_7C_8\log d}}{2C_7}\right)^2$, we have $C_7 t \geq \sqrt{t}+C_8\log d \geq \log t + C_8 \log d$. We can further bound $\mathbb{P}[B_i^c]$ by Assumption 1: $\mathbb{P}[B_i^c] = \mathbb{P}[\boldsymbol{x}_t^T(\boldsymbol{\beta}_j - \boldsymbol{\beta}_i) \leq 2\sqrt{s}\delta x_{\max}] \leq 2\sqrt{s} C_0 \delta x_{\max}$. Accordingly, we can bound the total regret of the final group as follows:

$$\begin{aligned}
&\sum_{t\geq (Kq)^2}^{T} \left\{4Kbx_{\max}\frac{C_2+1}{t} + 4C_0 s\delta^2 x_{\max}^2\right\} \\
\leq &\sum_{t\geq (Kq)^2}^{T} \left(4Kbx_{\max}\frac{C_2+1}{t} + \frac{C_0 s^2\sigma^2\lambda_{\max}}{2\lambda_{\min}^2 p^* t}\right) \\
\leq &\sum_{t\geq 4}^{T} \left(4Kbx_{\max}(C_2+1) + \frac{C_0 s^2\sigma^2\lambda_{\max}}{2\lambda_{\min}^2 p^*}\right)\frac{1}{t} \qquad (18)\\
\leq &\left(4Kbx_{\max}(C_2+1) + \frac{C_0 s^2\sigma^2\lambda_{\max}}{2\lambda_{\min}^2 p^*}\right)\log T, \qquad (19)
\end{aligned}$$

where Equation (18) requires $K \geq 2$, $q \geq 1$, and Equation (19) further requires $\sum_{t=4}^{T} 1/t \leq \log T$. Combining the bounds for these three groups, we can write the total regret as in Theorem 1.

Finally, for the dependence on d and s, we only need to consider the part with term $\log T$. As all $C_i,\ i = 2, 6$ are independence on d, $q = O(s^2 \log d)$ and $\lambda_{\max} \leq s x_{\max}^2$, we can derive that the regret dependence on d and s is upper bounded by $O\big(s^2(s + \log d)\big)$. □

Proof of Proposition 3

There are three key steps to prove Proposition 3:

1. Re-establish the compatibility condition for non-i.i.d. samples;
2. Provide the probability bound for the minimum eigenvalue of $\lambda_{\min} \doteq \frac{1}{|\mathcal{A}|}(\boldsymbol{X}_S^{\mathcal{A}})^T \boldsymbol{X}_S^{\mathcal{A}}$;
3. Refine the results in Proposition 1

For the first step, we will rely on the matrix perturbation techniques. As $|x_{t,i}|$ is bounded by $x_{\max}$ for all t and $i \leq d$, $\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}_S^{\mathcal{A}'})^T \boldsymbol{X}_S^{\mathcal{A}'}$ will converge to $\mathbb{E}[\boldsymbol{x}_S^T \boldsymbol{x}_S]$ when sample size is large enough. Similarly, for large sample size, the sample compatibility constant $\phi_{\mathcal{A}'}$ will also converge to the population compatibility constant ϕ. Based on this idea, the bound for $\phi_{\mathcal{A}}$ can be derived from ϕ with the bridge $\phi_{\mathcal{A}'}$. We summarize our results in Lemma 2.

Lemma 2 When $|\mathcal{A}'| \geq \frac{1024 s x_{\max}^3 \log d}{\phi^2}$, we will have $\phi_{|\mathcal{A}|} \geq \phi\sqrt{\frac{|\mathcal{A}'|}{2|\mathcal{A}|}}$ with probability $1 - \exp\left(-\frac{\phi^2|\mathcal{A}|}{512 s x_{\max}^2}\right)$.

In addition, we can show that $\mathrm{eig}_{\min}(\boldsymbol{A}+\boldsymbol{B}) > \mathrm{eig}_{\min}(\boldsymbol{A})$ if $\boldsymbol{A}, \boldsymbol{B} \succeq 0$. Immediately, we will have

$$\mathrm{eig}_{\min}\left(\frac{1}{|\mathcal{A}|}(\boldsymbol{X}_S^{\mathcal{A}})^T \boldsymbol{X}_S^{\mathcal{A}}\right) = \mathrm{eig}_{\min}\left(\frac{|\mathcal{A}'|}{|\mathcal{A}|}\frac{1}{|\mathcal{A}|}(\boldsymbol{X}_S^{\mathcal{A}'})^T \boldsymbol{X}_S^{\mathcal{A}'} + \frac{1}{|\mathcal{A}|}\sum_{i\in\mathcal{A}-\mathcal{A}'} \boldsymbol{x}_i^T \boldsymbol{x}_i\right) \geq \frac{|\mathcal{A}'|}{|\mathcal{A}|}\lambda_{\min,\mathcal{A}'}.$$

Combined with the matrix Chernoff bound ,the probability bound for $\lambda_{\min}$ can be established as in the following Lemma:

Lemma 3 (Tropp et al. [2015] Theorem 5.1.1) $\lambda_{\min,\mathcal{A}} \geq \frac{\lambda_{\min}|\mathcal{A}'|}{2|\mathcal{A}|}$ with probability $1 - s \cdot \exp\left(-\frac{|\mathcal{A}'|\lambda_{\min}}{8 x_{\max}^2}\right)$.

Note that Lemma 2 suggests $\phi_{\mathcal{A}} \geq \phi\sqrt{\frac{|\mathcal{A}'|}{2|\mathcal{A}|}} \geq \phi\sqrt{\frac{c_0}{4}}$ and Lemma 3 indicates $\mathrm{eig}_{\min}(\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}_S^{\mathcal{A}'})^T\boldsymbol{X}_S^{\mathcal{A}'}) = \lambda_{\min,\mathcal{A}} \geq \frac{\lambda_{\min}|\mathcal{A}'|}{2|\mathcal{A}|} \geq \frac{c_0}{4}\lambda_{\min}$. Plugging these two results back into Proposition 1, we can establish Proposition 3. □

Proof of Lemma 2

As all samples in $\mathcal{A}'$ are i.i.d, we will have the following result (see exercise 14.3 in Bühlmann & Van De Geer [2011]): if there exist K and σ_0 such that $K^2\left(E[\exp(x_{t,i}^2/K^2)-1]\right) \leq \sigma_0^2$, then

$$P\left\{\|\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}'} - E[\boldsymbol{x}^T\boldsymbol{x}]\|_\infty \geq 2K^2t + 2K\sigma_0\sqrt{2t} + 2K\sigma_0\lambda\left(\frac{K}{\sigma_0}, n, \binom{d}{2}\right)\right\} \leq \exp\left(-nt\right),$$

where $\lambda\left(\frac{K}{\sigma_0}, n, \binom{d}{2}\right) = \sqrt{\frac{2\log(d(d-1))}{n}} + \frac{K\log(d(d-1))}{n}$.

Setting $K = x_{\max}$ and $\sigma_0 = \sqrt{2}x_{\max}$, we will have $K^2\left(E[\exp(x_{t,i}^2/K^2)-1]\right) \leq x_{\max}^2(e-1) \leq \sigma_0^2$. Therefore, when the sample size is large enough, $\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}'}$ will not be far away from $E[\boldsymbol{x}^T\boldsymbol{x}]$ element-wise with high probability.

Now, we only need to show that if $\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}'}$ is close enough to $E[\boldsymbol{x}^T\boldsymbol{x}]$ element-wise, $\boldsymbol{X}^{\mathcal{A}'}$ will also satisfy the compatibility condition. To this end, we need Corollary 6.8 in Bühlmann & Van De Geer [2011], which shows that if 1) the population covariance matrix $E[\boldsymbol{x}^T\boldsymbol{x}]$ satisfies the compatibility condition with constant ϕ and 2) the sample covariance matrix $\frac{1}{n}\boldsymbol{X}^T\boldsymbol{X}$ satisfies $\|\frac{1}{n}\boldsymbol{X}^T\boldsymbol{X} - E[\boldsymbol{x}^T\boldsymbol{x}]\|_\infty \leq \frac{\phi^2}{2^5 s}$, then $\boldsymbol{X}$ will also satisfies the compatibility condition with constant $\phi_{\mathcal{A}} = \phi/\sqrt{2}$.

To combine these two results, we need to choose proper t and n so that the following inequality holds:

$$2K^2t + 2K\sigma_0\sqrt{2t} + 2K\sigma_0\lambda\left(\frac{K}{\sigma_0}, n, \binom{p}{2}\right) \leq \frac{\phi^2}{2^5 s}.$$

Intuitively, when t is sufficiently small and n is sufficiently large, the above inequality will hold naturally. Although there are various choices of t and n, we adopt the following parameter choice to simplify our proof: $t = \frac{\phi^2}{2^9 s x_{\max}^2}$ and $n \geq \frac{2^{10} s x_{\max}^3 \log d}{\phi^2}$, under which we will have

$$P\left\{\|\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}'} - E[\boldsymbol{x}^T\boldsymbol{x}]\|_\infty \leq \frac{\phi^2}{2^5 s}\right\} \geq 1 - \delta_4,$$

where $\delta_4 = \exp\left(-\frac{\phi^2}{2^9 s x_{\max}^2}\right)$. This result implies that $\phi_{\mathcal{A}'} = \frac{\phi}{\sqrt{2}}$ with probability $1-\delta_4$. The final step is to build the relationship between $\phi_{\mathcal{A}}$ and $\phi_{\mathcal{A}'}$. Consider a vector $\boldsymbol{\beta}$ with $\|\boldsymbol{\beta}_{S^c}\|_1 \leq 3\|\boldsymbol{\beta}_S\|_1$,

$$\begin{aligned}
\frac{1}{|\mathcal{A}|}\|\boldsymbol{X}^{\mathcal{A}}\boldsymbol{\beta}\|_2^2 &= \frac{1}{|\mathcal{A}|}\boldsymbol{\beta}^T(\boldsymbol{X}^{\mathcal{A}})^T\boldsymbol{X}^{\mathcal{A}}\boldsymbol{\beta} \\
&= \frac{1}{|\mathcal{A}|}\boldsymbol{\beta}^T\left((\boldsymbol{X}^{\mathcal{A}'})^T \quad (\boldsymbol{X}^{\mathcal{A}-\mathcal{A}'})^T\right)\begin{pmatrix}\boldsymbol{X}^{\mathcal{A}'} \\ \boldsymbol{X}^{\mathcal{A}-\mathcal{A}'}\end{pmatrix}\boldsymbol{\beta} \\
&= \frac{|\mathcal{A}'|}{|\mathcal{A}|}\boldsymbol{\beta}^T(\frac{1}{|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}'})\boldsymbol{\beta} + \frac{|\mathcal{A}|-|\mathcal{A}'|}{|\mathcal{A}|}\boldsymbol{\beta}^T(\frac{1}{|\mathcal{A}|-|\mathcal{A}'|}(\boldsymbol{X}^{\mathcal{A}-\mathcal{A}'})^T\boldsymbol{X}^{\mathcal{A}-\mathcal{A}'})\boldsymbol{\beta} \\
&\geq \frac{|\mathcal{A}'|}{|\mathcal{A}|}(\frac{1}{|\mathcal{A}'|}\|\boldsymbol{X}^{\mathcal{A}'}\boldsymbol{\beta}\|_2^2) \geq \frac{|\mathcal{A}'|}{|\mathcal{A}|}\frac{\phi_{\mathcal{A}'}^2}{s}\|\boldsymbol{\beta}\|_1^2 = \frac{|\mathcal{A}'|}{|\mathcal{A}|}\frac{\phi^2}{2s}\|\boldsymbol{\beta}\|_1^2,
\end{aligned}$$

where the last equality comes from the fact that $\phi_{\mathcal{A}'} = \frac{\phi}{\sqrt{2}}$. Therefore $\phi_{\mathcal{A}}$ is at least $\phi\sqrt{\frac{|\mathcal{A}'|}{2|\mathcal{A}|}}$.

□

Proof of Proposition 4 See Proposition 3 of Bastani & Bayati [2015]. $\square$

Proof of Proposition 5 For $x_t \in U_k$ where $k \in \mathcal{K}$, if the event $A_{t-1} = \left\{\|\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{i,t-1},\lambda) - \boldsymbol{\beta}^{true}\|_1 \le \frac{h}{4x_{\max}}\right\}$ holds, then we have:

$$\begin{aligned}
&\boldsymbol{x}_t^T\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{k,t-1},\lambda) - \boldsymbol{x}_t^T\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{i,t-1},\lambda)\\
\ge&\boldsymbol{x}_t^T(\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{k,t-1},\lambda) - \boldsymbol{\beta}_k^{true}) - \boldsymbol{x}_t^T(\hat{\boldsymbol{\beta}}_M(\mathcal{T}_{i,t-1},\lambda) - \boldsymbol{\beta}_i^{true}) + h\\
\ge& -\frac{1}{4}h - \frac{1}{4}h + h = \frac{1}{2}h.
\end{aligned}$$

Thus the set $\hat{K} = \{k|\boldsymbol{x}_t^T\hat{\boldsymbol{\beta}_M}(\mathcal{T}_{k,t-1},\lambda_1) \ge \max_{j\in K}\{\boldsymbol{x}_t^T\hat{\boldsymbol{\beta}_M}(\mathcal{T}_{j,t-1},\lambda_1)\} - h/2\}$ must be a singleton. As we use $\hat{K}$ as a pre-selection procedure in the MCP-Bandit algorithm, we will successfully select arm k for $x_t \in U_k$ if $\mathcal{A}_{t-1}$ holds.

Let's consider the probability of event A_{t-1} hold. When $T \ge 4$ and $q \ge \frac{8C_1 s\lambda_{\max}\sigma^2 x_{\max}}{h^2\lambda_{\min}}$ for the positive constant C_1, we have

$$\begin{aligned}
&\frac{\lambda_{\max}\sigma^2}{\lambda_{\min}}\cdot\frac{8sx_{\max}^2}{h^2} \le \frac{1}{C_1}q\log T \le |\mathcal{T}_{t,k}|\\
\Rightarrow&\sqrt{\frac{\lambda_{\max}\sigma^2}{2\lambda_{\min}}}\sqrt{\frac{s}{|\mathcal{A}|}} \le \frac{h}{4x_{\max}}.
\end{aligned}$$

From Proposition 3, we can show that there exists two positive constants C_4 and C_5 such that:

$$\mathbb{P}\{A_{t-1}\} \ge 1 - K\exp\left(-\frac{C_4}{C_1}q\log(t-1) + C_5\log d\right). \tag{20}$$

Furthermore, if we require $q \gtrsim O(s^2\log d)$, then we will have $q \ge \frac{C_1}{C_4}(C_5\log d + 1)$ for the constants C_1 and C_4, which implies the following result:

$$\mathbb{P}\{A_{t-1}\} \ge 1 - K\exp\left(-\frac{C_4}{C_1}q\log(t-1) + C_5\log d\right) = 1 - \frac{K}{t-1}. \tag{21}$$

Let's consider a sequence of $\{M(i), i = 0,1,2,..,T+1\}$ where $\{M(i)$ is defined as follows:

$$M(i) = \mathbb{E}\left[\sum_{j=1}^{T}\mathbf{1}(x_j\in U_k, A_{j-1}, j\notin\mathcal{T}_{T,k}|\mathcal{F}_i\right], \tag{22}$$

where $\mathcal{F}_i = \{\boldsymbol{x}_j, y_j, j\le i\}$. Thus $\{M(i)\}$ is a martingale with bounded difference $|M(i) - M(i+1)| \le 1, i = 0,1,2,...,T$. We can use $M(0)$ to bound the value of $M(T+1)$ with Azuma's inequality and we have:

$$\begin{aligned}
&\mathbb{P}(|M(T) - M(0)| \ge \frac{1}{2}M(0)) \le \exp\left(\frac{-M(0)^2/4}{2T}\right)\\
\Rightarrow\ &\mathbb{P}(M(0) - M(T) \ge \frac{1}{2}M(0)) \le \exp\left(\frac{-M(0)^2/4}{2T}\right)\\
\Rightarrow\ &\mathbb{P}(M(T) \le \frac{1}{2}M(0)) \le \exp\left(\frac{-M(0)^2/4}{2T}\right).
\end{aligned}$$

Further, we can bound $M(0)$ is bounded as follows:

$$\begin{aligned}
M(0) &= \mathbb{E}\left[\sum_{j=1}^{T}\mathbf{1}(x_j\in U_k, A_{j-1}, j\notin\mathcal{T}_{T,k})\right]\\
&= \sum_{j=1}^{T}\mathbb{E}[\mathbf{1}((x_j\in U_k, A_{j-1}, j\notin\mathcal{T}_{T,k})]\\
&= \sum_{j=1}^{T}\mathbb{P}((x_j\in U_k, A_{j-1}, j\notin\mathcal{T}_{T,k}).
\end{aligned} \tag{23}$$

Since $\{x \in U_i\}$ is independent on $\{A_{j-1}, j \notin \mathcal{T}_{T,k}\}$ and $\{j \notin \mathcal{T}_{T,k}\}$ is independent on $\{A_{j-1}\}$, Equation (23) implies:

$$\begin{aligned} M(0) &= \sum_{j=1}^{T} \mathbb{P}(x_j \in U_k)\mathbb{P}(A_{j-1})\mathbb{P}(j \notin \mathcal{T}_{T,k}) \\ &\geq p^*(1 - \frac{K}{T-1})(T - C_1 q \log T), \end{aligned} \tag{24}$$

where Equation (24) comes from Assumption 2, Equation (21), and the fact that the sample size of the forced samples is bounded for the constant C_1: $\frac{1}{C_1} q \log T \leq |\mathcal{T}_{t,k}| \leq C_1 q \log T$. When $T \geq \max\{(Kq)^2, (4C_1 q)^2\}$, we have

$$\begin{aligned} 1 - \frac{K}{T-1} \geq 1 - \frac{1}{Kq} \geq \frac{1}{2}, \\ (T - C_1 q \log T) \geq \frac{1}{2}T + (\frac{1}{2}T - C_1 q \sqrt{T}) \geq \frac{1}{2}T. \end{aligned}$$

Therefore, $M(0)$ is lower bounded by $\frac{p^*}{4}T$ and we have:

$$\begin{aligned} &\mathbb{P}\left(M(T) \leq \frac{p^* T}{8}\right) \leq \mathbb{P}\left(M(T) \leq M(0)\right) \leq \exp\left(\frac{-(p*)^2 T^2/64}{2T}\right) \\ \Rightarrow &\mathbb{P}\left(M(T) \leq \frac{p^* T}{8}\right) \leq \exp\left(-\frac{(p*)^2 T}{128}\right). \end{aligned} \tag{25}$$

Combining Equation (21) and Equation (25), we can conclude that with probability $1 - \exp\left((p^*)^2 T/128\right) - \frac{K}{T}$, among the whole sample set of arm $k \in \mathcal{K}$, the iid samples in U_k will be at least $p^* T/8$. The remaining proof directly follows the Proposition 3 with $c_0 = p^*/4$.

□

Reference

Bastani, H. and Bayati, M. Online decision-making with high-dimensional covariates. 2015.

Buhlmann, P. and Van De Geer, S. Statistics for high-dimensional data: methods, theory and applications. *Springer Science & Business Media*, 2011.

Fan, J., Xue, L., and Zou, H. Strong oracle optimality of folded concave penalized estimation. *Annals of statistics*, 42(3):819, 2014.

Rudelson, M., Vershynin, R., et al. Hanson-wright inequality and sub-gaussian concentration. *Electron. Commun. Probab*, 18(82):1-9, 2013.

Tropp, J. A. et al. An introduction to matrix concentration inequalities. *Foundations and Trends R in Machine Learning*, 8(1-2):1.

Van De Geer, S. A., Buhlmann, P., et al. On the conditions used to prove oracle results for the lasso. *Electronic Journal of Statistics*, 3:1360-1392, 2009.