Electronic Companion to "Online Learning and Decision-Making under Generalized Linear Model with High-Dimensional Data" by Xue Wang, Mike Mingcheng Wei, Tao Yao

Parameters	Explanation
t and T	Time indexes.
$\mathcal{K}$	The decision set: $\mathcal{K}=\{1,2,...,K\}$.
$R_{k,t}$ and R_i	The reward, where $k\in\mathcal{K}$, $t=1,2,...,T$, and $i=1,2,...,n$.
$\boldsymbol{X}_t$, $\boldsymbol{x}$, $\boldsymbol{x}_t$	The covariates vectors, where $\boldsymbol{X}_t,\boldsymbol{x},\boldsymbol{x}_t\in\mathbb{R}^d$, and $t=1,2,...,T$.
$d,\ s$	The dimension of total covariates and the dimension of significant covariates.
$\boldsymbol{\beta}_k^{\text{true}}$	User's true parameter vector corresponding to arm/decision k.
$f(\cdot),\ \mathcal{L}(\cdot)$	The sample-wise loss function and the negative log-likelihood loss function.
$f_y'(\cdot\|y),\ f_{yy}''(\cdot\|y)$	The first and second order partial derivatives of $f(\cdot\|y)$ with respect to y.
$x_{\max},\ R_{\max},\ b$	Positive constants that bound parameters defined in assumption A.1.
C	A positive constant defined in assumption A.2.
$\mathcal{K}_o,\ \mathcal{K}_s$	The optimal and suboptimal decision sets defined in assumption A.3.
U_k	A subset of users' covariates defined in assumption A.3, where $k\in\mathcal{K}$.
$h,\ p^*$	Positive constants defined in assumption A.3.
$\sigma,\ \sigma_2$	Positive constants defined in assumption A.4.
κ	The restricted eigenvalue constant defined in assumption A.5.
$\boldsymbol{\beta}^{\text{oracle}},\ \boldsymbol{\beta}^{\text{lasso}},\ \boldsymbol{\beta}^{\text{W}}$	The oracle, Lasso, and weighted Lasso estimators.
$\boldsymbol{\pi}$	The decision-makers' policy: $\boldsymbol{\pi}=\{\pi_t\}_{t\geq 1}$, where $\pi_t\in\mathcal{K}$ is the decision prescribed by policy $\boldsymbol{\pi}$ at time t.
$R^C(T)$	The cumulative regret up to time T.
$\mathcal{A}$	The sample set that contains only i.i.d. samples out of the whole sample set.
$\boldsymbol{w}$	Non-negative weights vector for weighted Lasso in Eq. (6), $\boldsymbol{w}=(w_1,w_2,...,w_d)$.
$P_{\lambda,a}(x)$	The MCP penalty function with positive parameters a and λ.
$\boldsymbol{\beta}^{\text{MCP}},\boldsymbol{\beta}^{\text{random}},\boldsymbol{\beta}^{\text{whole}}$	The MCP estimator, the MCP estimator under the random sample set $\mathcal{R}$, and the MCP estimator under the whole sample set $\mathcal{W}$.
$a,\lambda_1,\lambda_{2,0},t_0$	Input parameters for the G-MCP-Bandit algorithm.
$\delta_0(t,t_0)$	$\delta_0(t,t_0)\doteq 2((t_0+1)/(e(t+1)))^4$.
$\delta_1(n,\|\mathcal{A}\|,\zeta)$	$\delta_1(n,\|\mathcal{A}\|,\zeta)\doteq 2s\exp(-\frac{n\zeta^2}{2\sigma^2x_{\max}^2})+\exp(-C_1\|\mathcal{A}\|)$.
$\delta_2(n,\|\mathcal{A}\|,\lambda)$	$\delta_2(n,\|\mathcal{A}\|,\lambda)\doteq 4d\exp(-\frac{n\lambda^2}{2\sigma^2x_{\max}^2}\cdot(\frac{1}{2}-\frac{18ns}{\|\mathcal{A}\|\kappa a})^2)$.
$\mathcal{S}_1,\ \rho_{\mathcal{S}/\mathcal{S}_1}^{\text{MCP}}$	Terms defined for Proposition 1: $\mathcal{S}_1\doteq\{i:\|\beta_i^{\text{true}}\|\geq(\frac{24ns}{\|\mathcal{A}\|\kappa}+a)\lambda\}$; $\rho_{\mathcal{S}/\mathcal{S}_1}^{\text{MCP}}\doteq\\|\boldsymbol{\beta}_{\mathcal{S}/\mathcal{S}_1}^{\text{MCP}}-\boldsymbol{\beta}_{\mathcal{S}/\mathcal{S}_1}^{\text{true}}\\|_1/\\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{MCP}}-\boldsymbol{\beta}_{\mathcal{S}}^{\text{true}}\\|_1$ if $\mathcal{S}_1\neq\mathcal{S}$ and 0 otherwise.
$\mathcal{R}_{\boldsymbol{x},k}$	The set contains the user covariate $\boldsymbol{X}$ generated by the ϵ-decay random sampling method for arm k.
$\mathcal{F}_i$	A filtration defined as $\mathcal{F}_i=\{(\boldsymbol{X}_j,R_j)$ for $j\leq i\}$.
C_0	Defined in the proof of Proposition 4 and used in Proposition 2, 4-6, and Theorem 1; its dependence on T, d, and s is $C_0=\mathcal{O}(s^2\log d)$.
C_1	Defined in the statements of Lemma EC.1 and EC.2; $C_1=\mathcal{O}(s^{-2})$.
T_0,T_1	Defined in the proof of Proposition 6 and used in Proposition 6 and Theorem 1; $T_0=\tilde{\mathcal{O}}(s^2\log d)$ and $T_1=\tilde{\mathcal{O}}(\beta_{\min}^{-2}\cdot s^2\log d)$, where $\beta_{\min}=\min_{i\in\mathcal{S}^k,k\in\mathcal{K}}\|\beta_{k,i}^{\text{true}}\|$.
C_ρ,C_3,C_4,C_5	Defined in the proof of Theorem 1; $C_\rho=\mathcal{O}(1),\ C_3\leq\tilde{\mathcal{O}}((1+\rho_{\max})^2s^2\log d)$, $C_4=\mathcal{O}(1)$, and $C_5=\mathcal{O}(s^2)$.
$\mathcal{E}_0,\mathcal{E}_1,...,\mathcal{E}_{5,(i,j),t}(w)$	Series of events.
$\{M_k(i)\}$	Martingale sequences used in the proof of Proposition 5.

EC.1. Appendix: Main Proofs

To simplify the notation in the E-companion, we denote $\nabla_{\mathcal{B}} F(\boldsymbol{x})$ as the vector with elements $(\nabla F(\boldsymbol{x}))_i,\ i \in \mathcal{B}$, where $(\cdot)_i$ is the i-th element in the vector. Similarly, we denote $\nabla^2_{\mathcal{B},\mathcal{C}} F(\boldsymbol{x})$ as the matrix with elements $(\nabla^2 F(\boldsymbol{x}))_{ij},\ i \in \mathcal{B}, j \in \mathcal{C}$, where $(\cdot)_{ij}$ is the element in i-th column and j-th row. To prove the main lemma, propositions, and theorems in this section, we need four additional technical lemmas (i.e., Lemma EC.1 to Lemma EC.4), whose statements and proofs are given in §EC.2 of this E-Companion. For notational convenience, we will omit parameters' subscripts corresponding to the choice of arms, as long as doing so will not cause any misinterpretation.

Proof of Lemma 1 From the optimality condition of Eq. (3) and $\boldsymbol{\beta}^{\text{oracle}}$ being the optimal solution, we know that

$$\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{oracle}}) = \mathbf{0}. \tag{EC.1}$$

Expanding $\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta})$ in (EC.1) at $\boldsymbol{\beta}^{\text{true}}$, we can show that via the mean value theorem, for some $\boldsymbol{\xi} \in \{\tau \boldsymbol{\beta}^{\text{oracle}} + (1-\tau)\boldsymbol{\beta}^{\text{true}}, \tau \in [0,1]\}$, the following result holds:

$$\begin{aligned}
\nabla^2_{\mathcal{S},\mathcal{S}} \mathcal{L}(\boldsymbol{\xi})(\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}) &= \mathbf{0} - \nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) \\
\Rightarrow (\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}})^\top \nabla^2_{\mathcal{S},\mathcal{S}} \mathcal{L}(\boldsymbol{\xi})(\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}) &= -(\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}})^\top \nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) \\
\Rightarrow \boldsymbol{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi}) \boldsymbol{u} &= -(\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}})^\top \nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}}),
\end{aligned} \tag{EC.2}$$

where in (EC.2) we denote $\boldsymbol{u} = \boldsymbol{\beta}^{\text{oracle}} - \boldsymbol{\beta}^{\text{true}}$ and use the fact $\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}^c} = \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c} = \mathbf{0}$ to expend the left-hand side to d dimensional space. By the definition of $\boldsymbol{\beta}^{\text{oracle}}$ and $\boldsymbol{\beta}^{\text{true}}$, it is direct to show that $\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 = 0 \le 3\|\boldsymbol{u}_{\mathcal{S}}\|_1$. From Lemma EC.1, we know that when $|\mathcal{A}| \ge C_1^{-1} \log s$, the following inequality holds with probability at least $1 - \exp(-C_1|\mathcal{A}|)$:

$$\frac{|\mathcal{A}|\kappa}{2ns} \|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \le \boldsymbol{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi}) \boldsymbol{u}. \tag{EC.3}$$

Combining (EC.2) and (EC.3), we have:

$$\begin{aligned}
\frac{|\mathcal{A}|\kappa}{2ns} \|\boldsymbol{u}_{\mathcal{S}}\|_1^2 &\le -(\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}})^\top \nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) \\
\Rightarrow \frac{|\mathcal{A}|\kappa}{2ns} \|\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1^2 &\le \|\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1 \|\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty \\
\Rightarrow \|\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1 &\le \frac{2ns}{|\mathcal{A}|\kappa} \|\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty.
\end{aligned} \tag{EC.4}$$

To obtain an upper bound for $\|\boldsymbol{\beta}^{\text{oracle}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1$, we need to show that $\|\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty$ is also upper bounded.

- **Upper bound for $\|\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty$:**

From the definition of $\mathcal{L}(\cdot)$, we have

$$\|\nabla_{\mathcal{S}} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty = \left\| \frac{1}{n} \sum_{j=1}^{n} (\boldsymbol{X}_{j,\mathcal{S}})^\top f_y'(R_j | \boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}}) \right\|_\infty, \tag{EC.5}$$

where we replace r and y in $f_y'(r|y)$ by R_j and $\boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}}$ respectively, and $\boldsymbol{X}_{j,\mathcal{S}}$ is the subvector of $\boldsymbol{X}_j$ with elements in $\mathcal{S}$. Under assumption A.4, $f_y'(R_j|\boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}})$ is a σ^2-sub-gaussian random variable. From the Hoeffding inequality (see Proposition 2.5 in Wainwright 2019), for $\zeta > 0$, we have

$$\mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}(X_{j,i})f_y'(R_j|\boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}})\right| \geq \zeta\right) \leq 2\exp\left(-\frac{n\zeta^2}{2\sigma^2 x_{\max}^2}\right) \quad \forall i \in \mathcal{S}, \tag{EC.6}$$

where the right-hand side uses the fact that all realization $\|\boldsymbol{x}_j\|_\infty \leq x_{\max}$ in assumption A.1. Hence, via union bound, we can show that

$$\begin{aligned}\mathbb{P}\left(\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty \geq \zeta\right) &= \mathbb{P}\left(\left\|\frac{1}{n}\sum_{j=1}^{n}(\boldsymbol{X}_{j,\mathcal{S}})^\top f_y'(R_j|\boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}})\right\|_\infty \geq \zeta\right)\\ &\leq \sum_{i\in\mathcal{S}}\mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}X_{j,i}f_y'(R_j|\boldsymbol{X}_j^\top \boldsymbol{\beta}^{\text{true}})\right| \geq \zeta\right)\\ &\leq 2s\exp\left(-\frac{n\zeta^2}{2\sigma^2 x_{\max}^2}\right),\end{aligned} \tag{EC.7}$$

where the last inequality in (EC.7) follows from $|\mathcal{S}| \leq s$. At last, the lemma follows directly by combining (EC.4) and (EC.7).

Proof of Proposition 1 Proposition 1 directly follows Proposition 3 by setting $|\mathcal{A}| = n$.

Proof of Proposition 2 Under the ϵ-decay random sampling method, the probability of randomly drawn arm k at time t is $\min\{1, t_0/t\}/|\mathcal{K}|$, where $|\mathcal{K}|$ is the number of arms. Hence, at time T, the expected total number of times at which arm k was randomly drawn is

$$\mathbb{E}[n_k] = \frac{1}{|\mathcal{K}|}\sum_{t=1}^{T}\min\left\{1, \frac{t_0}{t}\right\},$$

where the expectation is taken with respect to n_k, the total number of random samples.

When $T > t_0$,

$$\mathbb{E}[n_k] = \frac{1}{|\mathcal{K}|}\left(t_0 + \sum_{t=t_0+1}^{T}\frac{t_0}{t}\right) = \frac{t_0}{|\mathcal{K}|}\left(1 + \sum_{t=t_0+1}^{T}\frac{1}{t}\right). \tag{EC.8}$$

Since the function $f(t) = 1/t$ is decreasing in t, it can be bounded as follows

$$\int_t^{t+1}\frac{1}{x}dx < \frac{1}{t} < \int_{t-1}^{t}\frac{1}{x}dx,\ t \geq 2.$$

As $C_0 \geq 20$, we can verify that $t_0 = 2C_0|\mathcal{K}| \geq 2$. Hence, for any t from $t_0 + 1$ to T, we have

$$\log(T+1) - \log(t_0+1) < \sum_{t=t_0+1}^{T}\frac{1}{t} < \log(T) - \log(t_0). \tag{EC.9}$$

Combining (EC.8) and (EC.9), we can bound $\mathbb{E}[n_k]$ as follows:

$$\frac{1}{|\mathcal{K}|}t_0(1 + \log(T+1) - \log(t_0+1)) < \mathbb{E}[n_k] < \frac{1}{|\mathcal{K}|}t_0(1 + \log(T) - \log(t_0)). \tag{EC.10}$$

Since $n_k = \sum_{t=1}^{T}\mathbb{1}\{\text{random sampling for arm } k \text{ at time } t\}$, we can view n_k as the summation of bounded i.i.d. random variables. By Chernoff bound (see Theorem 4 in Goemans 2015 by setting $\delta = 0.5$), we can have the following inequality:

$$\mathbb{P}\left(\frac{1}{2}\mathbb{E}[n_k] \le n_k \le \frac{3}{2}\mathbb{E}[n_k]\right) \ge 1 - 2\exp\left(-\frac{1}{10}\mathbb{E}[n_k]\right). \tag{EC.11}$$

We then relax the $\mathbb{E}[n_k]$ in (EC.11) by using the upper and lower bounds provided in (EC.10) to attain the following result:

$$\mathbb{P}\left(\frac{t_0(1+\log(T+1)-\log(t_0+1))}{2|\mathcal{K}|} \le n_k \le \frac{3t_0(1+\log(T)-\log(t_0))}{2|\mathcal{K}|}\right) \quad \ge 1 - 2\left(\frac{t_0+1}{e(T+1)}\right)^{\frac{t_0}{10|\mathcal{K}|}}. \tag{EC.12}$$

When $t_0 = 2C_0|\mathcal{K}|$ and $C_0 \ge 20$, we have $\frac{t_0}{10|\mathcal{K}|} = C_0/5 \ge 4$. Then, this proposition follows directly by plugging $t_0 = 2C_0|\mathcal{K}|$ back into (EC.12) and using the definition of $\delta_0(T, t_0)$ in the proposition statement.

Proof of Proposition 3 In the first step of the 2sWL procedure, we solve a Lasso problem. From Lemma EC.2, we know that if $|\mathcal{A}| \ge C_1^{-1}\log d$, then the inequality $\|\boldsymbol{\beta}^{\text{lasso}} - \boldsymbol{\beta}^{\text{true}}\|_1 \le \frac{24ns\lambda}{|\mathcal{A}|\kappa}$ holds with high probability. Beside set $\mathcal{S}_1$ defined in (9), let's consider the following index set:

$$\mathcal{S}_2 \doteq \left\{i : |\beta_i^{\text{true}}| < \left(\frac{24ns}{|\mathcal{A}|\kappa} + a\right)\lambda, i \in \mathcal{S}\right\}. \tag{EC.13}$$

Directly, we can show that

$$\begin{aligned} &i \in \mathcal{S}_1 \Rightarrow |\beta_i^{\text{lasso}}| \ge a\lambda \text{ so that } w_i = P'_{\lambda,a}(|\beta_i^{\text{lasso}}|) = 0; \\ &i \in \mathcal{S}_2 \Rightarrow |\beta_i^{\text{lasso}}| \le \left(\frac{48ns}{|\mathcal{A}|\kappa} + a\right)\lambda \text{ and } w_i = P'_{\lambda,a}(|\beta_i^{\text{lasso}}|) \le \lambda, \end{aligned} \tag{EC.14}$$

where we use the fact that for all $x \ge 0$

$$P'_{\lambda,a}(x) = \max\left(0, \lambda - \frac{x}{a}\right) \tag{EC.15}$$

per definition of MCP penalty in (7). Similarly, for $i \in \mathcal{S}^{\text{c}} = \{i : |\beta_i^{\text{true}}| = 0, i \in \{1, 2, ..., d\}\}$, we can show that

$$i \in \mathcal{S}^{\text{c}} \Rightarrow |\beta_i^{\text{lasso}}| \le \frac{24ns}{|\mathcal{A}|\kappa}\lambda \text{ and } w_i = P'_{\lambda,a}(|\beta_i^{\text{lasso}}|) \ge \left(1 - \frac{24ns}{|\mathcal{A}|\kappa a}\right)\lambda, \tag{EC.16}$$

where the last inequality uses $1 - \frac{24ns}{|\mathcal{A}|\kappa a} > 0$ for $a > \frac{48ns}{|\mathcal{A}|\kappa}$.

Let $\boldsymbol{\beta}^{\text{MCP}}$ be the optimal solution to the second step of the 2sWL procedure. Using the fact that $\mathcal{L}(\boldsymbol{\beta}) + \sum_{j=1}^d w_j|\beta_j|$ is minimized at $\boldsymbol{\beta}^{\text{MCP}}$ and the fact that $\mathcal{L}(\boldsymbol{\beta})$ is convex, we have

$$\begin{aligned} &\mathcal{L}(\boldsymbol{\beta}^{\text{MCP}}) + \sum_{j=1}^d w_j|\beta_j^{\text{MCP}}| \le \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \sum_{j=1}^d w_j|\beta_j^{\text{true}}| && \text{(EC.17)} \\ \Rightarrow &\mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}) + \sum_{j=1}^d w_j|\beta_j^{\text{MCP}}| \le \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \sum_{j=1}^d w_j|\beta_j^{\text{true}}| \\ \Rightarrow &\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}) + \sum_{j=1}^d w_j|\beta_j^{\text{MCP}}| \le \sum_{j=1}^d w_j|\beta_j^{\text{true}}| \\ \Rightarrow &\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}) + \sum_{j\in\mathcal{S}_2} w_j|\beta_j^{\text{MCP}}| + \sum_{j\in\mathcal{S}^c} w_j|\beta_j^{\text{MCP}}| \le \sum_{j\in\mathcal{S}_2} w_j|\beta_j^{\text{true}}| && \text{(EC.18)} \\ \Rightarrow &\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}) + \sum_{j\in\mathcal{S}^c} w_j|\beta_j^{\text{MCP}} - \beta_j^{\text{true}}| \le \sum_{j\in\mathcal{S}_2} w_j|\beta_j^{\text{MCP}} - \beta_j^{\text{true}}|. && \text{(EC.19)} \end{aligned}$$

where (EC.18) uses the observations that $w_i = 0$ for $i \in \mathcal{S}_1$ and $\beta_i^{\text{true}} = 0$ for $i \in \mathcal{S}^{\text{c}}$, and (EC.19) uses the observation that $\beta_i^{\text{true}} = 0$ for $i \in \mathcal{S}^{\text{c}}$.

Let $\boldsymbol{u} = \boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}$. Then, inequality (EC.19) can be further simplified as follows:

$$\begin{aligned}
&\nabla \mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top \boldsymbol{u} + \sum_{j\in\mathcal{S}^c} w_j|u_j| \le \sum_{j\in\mathcal{S}_2} w_j|u_j| && \text{(EC.20)}\\
\Rightarrow &\sum_{j\in\mathcal{S}_1\cup\mathcal{S}_2\cup\mathcal{S}^c} \nabla_j \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) u_j + \sum_{j\in\mathcal{S}^c} w_j|u_j| \le \sum_{j\in\mathcal{S}_2} w_j|u_j| \\
\Rightarrow &\sum_{j\in\mathcal{S}^c} (w_j - |\nabla_j \mathcal{L}(\boldsymbol{\beta}^{\text{true}})|)|u_j| \le \sum_{j\in\mathcal{S}_2} (w_j + |\nabla_j \mathcal{L}(\boldsymbol{\beta}^{\text{true}})|)|u_j| + \sum_{j\in\mathcal{S}_1} |\nabla_j \mathcal{L}(\boldsymbol{\beta}^{\text{true}})||u_j| \\
\Rightarrow &(\tilde{w}_c - \|\nabla_{\mathcal{S}^c} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty)\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le (\tilde{w}_2 + \|\nabla_{\mathcal{S}_2} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty)\|\boldsymbol{u}_{\mathcal{S}_2}\|_1 + \|\nabla_{\mathcal{S}_1} \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty \|\boldsymbol{u}_{\mathcal{S}_1}\|_1, && \text{(EC.21)}
\end{aligned}$$

where we define two positive constants, $\tilde{w}_c$ and $\tilde{w}_2$, as follows:

$$\tilde{w}_c \doteq \left(1 - \frac{24ns}{|\mathcal{A}|\kappa a}\right)\lambda \le \min_{j\in\mathcal{S}^c}\{w_j\} \qquad \text{(EC.22)}$$

and

$$\tilde{w}_2 \doteq \lambda \ge \max_{j\in\mathcal{S}_2}\{w_j\}, \qquad \text{(EC.23)}$$

where the inequalities in (EC.22) and (EC.23) are from (EC.16) and (EC.14), respectively.

Now, we define the following event:

$$\mathcal{E}_{\text{sub},1} \doteq \left\{ \|\nabla \mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty < \frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2 \right\}. \qquad \text{(EC.24)}$$

Then, under event $\mathcal{E}_{\text{sub},1}$, inequality (EC.21) implies

$$\begin{aligned}
&\left(\tilde{w}_c - \frac{3}{4}\tilde{w}_c + \frac{1}{4}\tilde{w}_2\right)\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le \left(\tilde{w}_2 + \frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2\right)\|\boldsymbol{u}_{\mathcal{S}_2}\|_1 + \left(\frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2\right)\|\boldsymbol{u}_{\mathcal{S}_1}\|_1 \\
\Rightarrow &\frac{1}{4}(\tilde{w}_c + \tilde{w}_2)\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le \frac{3}{4}(\tilde{w}_2 + \tilde{w}_c)\|\boldsymbol{u}_{\mathcal{S}_2}\|_1 + \frac{3}{4}(\tilde{w}_2 + \tilde{w}_c)\|\boldsymbol{u}_{\mathcal{S}_1}\|_1 - \tilde{w}_2\|\boldsymbol{u}_{\mathcal{S}_1}\|_1 \\
\Rightarrow &(\tilde{w}_c + \tilde{w}_2)\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le 3(\tilde{w}_2 + \tilde{w}_c)\|\boldsymbol{u}_{\mathcal{S}}\|_1 - 4\tilde{w}_2\|\boldsymbol{u}_{\mathcal{S}_1}\|_1 \\
\Rightarrow &\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le 3\|\boldsymbol{u}_{\mathcal{S}}\|_1 - \frac{4\tilde{w}_2}{\tilde{w}_c + \tilde{w}_2}\|\boldsymbol{u}_{\mathcal{S}_1}\|_1 \\
\Rightarrow &\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \le 3\|\boldsymbol{u}_{\mathcal{S}}\|_1. && \text{(EC.25)}
\end{aligned}$$

Combining (EC.25) and Lemma EC.1, we can show that for all feasible $\boldsymbol{\xi}$, the following inequality holds:

$$\mathbb{P}\left(\frac{|\mathcal{A}|\kappa}{2ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \le u^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi})\boldsymbol{u}\right) \ge 1 - \exp(-C_1|\mathcal{A}|). \qquad \text{(EC.26)}$$

Now, we go back to (EC.17) and expand the $\mathcal{L}(\boldsymbol{\beta})$ term in the left-hand side at $\boldsymbol{\beta}^{\text{true}}$. Denoting $\boldsymbol{u} = \boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}$, we can show that there exists a feasible $\boldsymbol{\xi}$ between $\boldsymbol{\beta}^{MCP}$ and $\boldsymbol{\beta}^{true}$ such that

$$\begin{aligned}
&\mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \nabla \mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top \boldsymbol{u} + \frac{1}{2}\boldsymbol{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi})\boldsymbol{u} + \sum_{i=1}^d w_i|\beta_i^{\text{MCP}}| \le \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \sum_{i=1}^d w_i|\beta_i^{\text{true}}| \\
\Rightarrow &\nabla \mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top \boldsymbol{u} + \frac{1}{2}\boldsymbol{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi})\boldsymbol{u} + \sum_{i=1}^d w_i|\beta_i^{\text{MCP}}| \le \sum_{i=1}^d w_i|\beta_i^{\text{true}}| \\
\Rightarrow &\nabla \mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top \boldsymbol{u} + \frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 + \sum_{i=1}^d w_i|\beta_i^{\text{MCP}}| \le \sum_{i=1}^d w_i|\beta_i^{\text{true}}| && \text{(EC.27)}\\
\Rightarrow &\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \le \sum_{i=1}^d \left(-\nabla_i \mathcal{L}(\boldsymbol{\beta}^{\text{true}})u_i + w_i(|\beta_i^{\text{true}}| - |\beta_i^{\text{MCP}}|)\right)
\end{aligned}$$

$$\Rightarrow \frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \leq \sum_{i\in\mathcal{S}_1\cup\mathcal{S}_2\cup\mathcal{S}^c}\left(-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})u_i + w_i(|\beta_i^{\text{true}}| - |\beta_i^{\text{MCP}}|)\right), \tag{EC.28}$$

where inequality (EC.27) uses (EC.26) and we defer the consideration of the probability part via union bound to (EC.34). Then, we can bound the right hand side of (EC.28) by considering $i\in\mathcal{S}_1,\mathcal{S}_2$ and $\mathcal{S}^c$ separately.

• $i\in\mathcal{S}_1$**:**

$$\begin{aligned}
&\sum_{i\in\mathcal{S}_1}\left(-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})u_i + w_i(|\beta_i^{\text{true}}| - |\beta_i^{\text{MCP}}|)\right)\\
\leq&\sum_{i\in\mathcal{S}_1}\left(\left|\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right| + w_i\right)|u_i|\\
=&\sum_{i\in\mathcal{S}_1}\left(\left|\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right|\right)|u_i|\\
\leq&\|\boldsymbol{u}_{\mathcal{S}_1}\|_1\|\nabla_{\mathcal{S}_1}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty,
\end{aligned} \tag{EC.29}$$

where the equality uses $w_i=0$ for all $i\in\mathcal{S}_1$.

• $i\in\mathcal{S}_2$**:**

$$\begin{aligned}
&\sum_{i\in\mathcal{S}_2}\left(-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})u_i + w_i(|\beta_i^{\text{true}}| - |\beta_i^{\text{MCP}}|)\right)\\
\leq&\sum_{i\in\mathcal{S}_2}\left(\left|\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right| + w_i\right)|u_i|\\
\leq&\|\boldsymbol{u}_{\mathcal{S}_2}\|_1\left(\|\nabla_{\mathcal{S}_2}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \lambda\right),
\end{aligned} \tag{EC.30}$$

where the last inequality uses $w_i\leq\lambda$ for $i\in\mathcal{S}_2$.

• $i\in\mathcal{S}^c$**:**

$$\begin{aligned}
&\sum_{i\in\mathcal{S}^c}\left(-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})u_i + w_i(|\beta_i^{\text{true}}| - |\beta_i^{\text{MCP}}|)\right)\\
=&\sum_{i\in\mathcal{S}^c}\left(-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\beta_i^{\text{MCP}} - w_i|\beta_j^{\text{MCP}}|\right)\\
\leq&\sum_{i\in\mathcal{S}^c}\left(|\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})||\beta_i^{\text{MCP}}| - w_i|\beta_i^{\text{MCP}}|\right)\\
\leq&\sum_{i\in\mathcal{S}^c}\left(\frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2 - w_i\right)|\beta_i^{\text{MCP}}|\\
\leq&0,
\end{aligned} \tag{EC.31}$$

where in the second-to-last inequality, we use the event $\mathcal{E}_{\text{sub},1}$ in (EC.24), and in the last inequality, we adopt the fact that $\tilde{w}_c\leq w_i$, $\tilde{w}_2>0$, and $w_i>0$ by definitions.

Then, we combine (EC.28), (EC.29), (EC.30) and (EC.31):

$$\begin{aligned}
&\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \leq \|\boldsymbol{u}_{\mathcal{S}_1}\|_1\|\nabla_{\mathcal{S}_1}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \|\boldsymbol{u}_{\mathcal{S}_2}\|_1\left(\|\nabla_{\mathcal{S}_2}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \lambda\right)\\
\Rightarrow&\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \leq \|\boldsymbol{u}_{\mathcal{S}}\|_1\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \lambda\|\boldsymbol{u}_{\mathcal{S}_2}\|_1\\
\Rightarrow&\|\boldsymbol{u}_{\mathcal{S}}\|_1 \leq \frac{4ns}{|\mathcal{A}|\kappa}\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \frac{4ns}{|\mathcal{A}|\kappa}\cdot\lambda\cdot\frac{\|\boldsymbol{u}_{\mathcal{S}_2}\|_1}{\|\boldsymbol{u}_{\mathcal{S}}\|_1}\\
\Rightarrow&\|\boldsymbol{u}\|_1 = \|\boldsymbol{u}_{\mathcal{S}}\|_1 + \|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \leq 4\|\boldsymbol{u}_{\mathcal{S}}\|_1 \leq \frac{16ns}{|\mathcal{A}|\kappa}\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \frac{16ns}{|\mathcal{A}|\kappa}\cdot\lambda\cdot\frac{\|\boldsymbol{u}_{\mathcal{S}_2}\|_1}{\|\boldsymbol{u}_{\mathcal{S}}\|_1}
\end{aligned} \tag{EC.32}$$

$$\Rightarrow \|\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}\|_1 \le \frac{16ns}{|\mathcal{A}|\kappa}\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \frac{16ns}{|\mathcal{A}|\kappa} \cdot \frac{\|\boldsymbol{\beta}^{\text{MCP}}_{\mathcal{S}_2} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}_2}\|_1}{\|\boldsymbol{\beta}^{\text{MCP}}_{\mathcal{S}} - \boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1} \cdot \lambda$$
$$\Rightarrow \|\boldsymbol{\beta}^{\text{MCP}} - \boldsymbol{\beta}^{\text{true}}\|_1 \le \frac{16ns}{|\mathcal{A}|\kappa}\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty + \frac{16ns}{|\mathcal{A}|\kappa} \cdot \rho^{\text{MCP}}_{\mathcal{S}/\mathcal{S}_1} \cdot \lambda,$$

where first inequality in (EC.32) applies (EC.25), and we use the definition of $\rho^{\text{MCP}}_{\mathcal{S}/\mathcal{S}_1}$ (i.e., Equation (12)) in the last inequality. Then, via (EC.7), we can bound $\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty$ as follows:

$$\mathbb{P}(\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty \le \zeta) \ge 1 - 2s\exp\left(-\frac{n\zeta^2}{2\sigma^2 x_{\max}^2}\right). \tag{EC.33}$$

Next, we build the probability bound for event $\mathcal{E}_{\text{sub},1}$ in (EC.24) by the Hoeffding's inequality. For $t > 0$, from (EC.7), we have

$$\mathbb{P}\left(\left\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_\infty \ge t\right) \le 2d\exp\left(-\frac{nt^2}{2\sigma^2 x_{\max}^2}\right)$$
$$\Rightarrow \mathbb{P}\left(\left\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_\infty \ge \frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2\right) \le 2d\exp\left(-\frac{n(\frac{3}{4}\tilde{w}_c - \frac{1}{4}\tilde{w}_2)^2}{2\sigma^2 x_{\max}^2}\right)$$
$$\Rightarrow \mathbb{P}\left(\left\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_\infty \ge \left(\frac{1}{2} - \frac{18ns}{|\mathcal{A}|\kappa a}\right)\lambda\right) \le 2d\exp\left(-\frac{n\lambda^2}{2\sigma^2 x_{\max}^2} \cdot \left(\frac{1}{2} - \frac{18ns}{|\mathcal{A}|\kappa a}\right)^2\right).$$

Combining this result with Lemma EC.2, we can show that if $|\mathcal{A}| \ge C_1^{-1}\log d$ and $a > \frac{48ns}{|\mathcal{A}|\kappa}$ (which also implies that $\frac{1}{2} > \frac{18ns}{|\mathcal{A}|\kappa a}$), the proposition statement holds with probability

$$1 - \exp(-C_1|\mathcal{A}|) - 2d\exp\left(-\frac{n\lambda^2}{8\sigma^2 x_{\max}}\right) - 2s\exp\left(-\frac{n\zeta^2}{2\sigma^2 x_{\max}}\right) - 2d\exp\left(-\frac{n\lambda^2}{2\sigma^2 x_{\max}} \cdot \left(\frac{1}{2} - \frac{18ns}{|\mathcal{A}|\kappa a}\right)^2\right)$$
$$\ge 1 - \exp(-C_1|\mathcal{A}|) - 4d\exp\left(-\frac{n\lambda^2}{2\sigma^2 x_{\max}} \cdot \left(\frac{1}{2} - \frac{18ns}{|\mathcal{A}|\kappa a}\right)^2\right) - 2s\exp\left(-\frac{n\zeta^2}{2\sigma^2 x_{\max}^2}\right), \tag{EC.34}$$

where the last inequality uses the fact that $\left(\frac{1}{2} - \frac{18ns}{|\mathcal{A}|\kappa a}\right)^2 \le \frac{1}{4}$.

Proof of Proposition 4 For clear expositions, we first state two constants that we will use in this proof:

$$C_0 \ge \max\left\{20, \frac{64}{p^*}, \frac{24\log d}{p^* C_1}, \frac{96}{p^* C_1}, \frac{3072\sigma^2 x_{\max}^2(1+\log d)}{\lambda^2}\right\} \tag{EC.35}$$

and

$$\lambda \le \min\left\{\frac{h\kappa p^*}{3072 e\sigma s R_{\max} x_{\max}}, \frac{p^*\kappa}{768\sigma s x_{\max}}\right\}. \tag{EC.36}$$

As $C_1 = \mathcal{O}(s^{-2})$ per equation (EC.125), it is direct to verify that $C_0 = \mathcal{O}(s^2\log d)$. Further, we denote event $\mathcal{E}_3$ as follows:

$$\mathcal{E}_3 = \left\{\frac{|\mathcal{A}|}{n} \ge \frac{1}{24}p^*\right\}. \tag{EC.37}$$

Note that for the suboptimal arm set (i.e., $k \in \mathcal{K}_s$), event $\mathcal{E}_3$ holds automatically, as $|\mathcal{A}| = n$ for $k \in \mathcal{K}_s$ so that $|\mathcal{A}|/n > p^*/24$ always holds true; for the optimal arm set (i.e., $k \in \mathcal{K}_o$), $|\mathcal{A}|/n$ represents the proportion of covariate vectors $\boldsymbol{X}$ that are in the set U_k (i.e., $\boldsymbol{X} \in U_k$) to all n i.i.d. samples that are generated via the ϵ-decay sampling scheme, and we will bound the probability for event $\mathcal{E}_3$ later in (EC.45).

Combining Proposition 2 and event $\mathcal{E}_3$, we can show that when $C_0 \geq \max\left\{20, \frac{24\log d}{p^* C_1}\right\}$, the following inequalities hold simultaneously with probability at least $1-\delta_0(t,t_0)$:

$$|\mathcal{A}| \geq \frac{p^*}{24} n \geq \frac{p^*}{24} C_0(1+\log(t+1)-\log(t_0+1)) \geq \frac{p^*}{24} C_0 \geq C_1^{-1}\log d. \tag{EC.38}$$

In addition, given $\mathcal{E}_3$, we can show that under the condition $a > \frac{1152s}{p^*\kappa}$, the following inequality holds:

$$a > \frac{1152s}{p^*\kappa} \geq \frac{48ns}{\kappa|\mathcal{A}|}. \tag{EC.39}$$

Hence, with (EC.38) and (EC.39), it is direct to show that the inequality (16) in Proposition 3 holds: for $\zeta > 0$, we have the following inequality:

$$\begin{aligned}
&\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{random}} - \boldsymbol{\beta}^{\text{true}}\|_1 \leq \frac{16ns\zeta}{|\mathcal{A}|\kappa} + \frac{16ns\rho^{\text{random}}_{\mathcal{S}/\mathcal{S}_1}}{|\mathcal{A}|\kappa}\lambda\right) \geq 1-\delta_1(n,|\mathcal{A}|,\zeta) - \delta_2(n,|\mathcal{A}|,\lambda) \\
\Rightarrow &\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{random}} - \boldsymbol{\beta}^{\text{true}}\|_1 \leq \frac{32ns}{|\mathcal{A}|\kappa}\lambda\right) \geq 1-\delta_1(n,|\mathcal{A}|,\lambda) - \delta_2(n,|\mathcal{A}|,\lambda),
\end{aligned} \tag{EC.40}$$

where in (EC.40), we set $\zeta = \lambda$ and use $\rho^{\text{random}}_{\mathcal{S}/\mathcal{S}_1} \leq 1$.

Combining event $\mathcal{E}_3$ and Proposition 2, we can show that with probability at least $1-\delta_0(t,t_0)$, the following results hold

$$n \geq C_0(1+\log(t+1)-\log(t_0+1)) \text{ and } |\mathcal{A}| \geq \frac{p^*}{24} n \geq \frac{p^*}{24} C_0(1+\log(t+1)-\log(t_0+1)). \tag{EC.41}$$

Then, we can further simplify (EC.41) as follows:

$$n \geq C_0 \log\left(\frac{e(t+1)}{t_0+1}\right) \text{ and } |\mathcal{A}| \geq \frac{p^*}{24} C_0 \log\left(\frac{e(t+1)}{t_0+1}\right). \tag{EC.42}$$

Now, if we set $C_0 \geq \max\left\{\frac{96}{p^* C_1}, \frac{3072\sigma^2 x_{\max}^2(1+\log d)}{\lambda^2}\right\}$, then we can directly verify that $\delta_1(n,|\mathcal{A}|,\lambda) = 2s\exp\left(-\frac{n\lambda^2}{2\sigma^2 x_{\max}^2}\right) + \exp(-C_1|\mathcal{A}|) \leq \delta_0(t,t_0) + \frac{1}{2}\delta_0(t,t_0)$ and $\delta_2(n,|\mathcal{A}|,\lambda) \leq 2\delta_0(t,t_0)$, combining which we have the following result:

$$\delta_1(n,|\mathcal{A}|,\lambda) + \delta_2(n,|\mathcal{A}|,\lambda) \leq \frac{7}{2}\delta_0(t,t_0). \tag{EC.43}$$

Next, we need to bound the probability for event $\mathcal{E}_3$ for $k \in \mathcal{K}_o$. First, we can show that

$$\begin{aligned}
\left\{\frac{|\mathcal{A}|}{n} \geq \frac{1}{24}p^*\right\} &\supseteq \left\{|\mathcal{A}| \geq \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} \cap \left\{n \leq 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} \\
&= \left(\left\{|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} \cup \left\{n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}\right)^{\mathrm{c}},
\end{aligned} \tag{EC.44}$$

which infers that for $k \in \mathcal{K}_o$,

$$\begin{aligned}
\mathbb{P}\left\{\frac{|\mathcal{A}|}{n} \geq \frac{1}{24}p^*\right\} &\geq \mathbb{P}\left\{\left(\left\{|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} \cup \left\{n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}\right)^{\mathrm{c}}\right\} \\
&= 1 - \mathbb{P}\left\{\left\{|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} \cup \left\{n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}\right\} \\
&\geq 1 - \mathbb{P}\left\{|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\} - \mathbb{P}\left\{n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}.
\end{aligned} \tag{EC.45}$$

Now, we will separately consider bounds for $\mathbb{P}\left\{|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}$ and $\mathbb{P}\left\{n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right\}$.

• **The probability bound for** $n > 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)$**:**

From Proposition 2, when $t \geq t_0$, the following result holds with probability $1-\delta_0(t,t_0)$:

$$\begin{aligned} n \leq 3C_0(1+\log(t)-\log(t_0)) &= 3C_0 \log\left(\frac{et}{t_0}\right) \\ &< 3C_0 \log\left(\frac{2e(t+1)}{2t_0}\right) \\ &< 3C_0 \log\left(\frac{2e(t+1)}{t_0+1}\right) \\ &= 3C_0 \log\left(\frac{e(t+1)}{t_0+1}\right) + 3C_0 \log(2) \\ &< 6C_0 \log\left(\frac{e(t+1)}{t_0+1}\right), \end{aligned} \tag{EC.46}$$

where the last inequality uses $2 < e < \frac{e(t+1)}{t_0+1}$.

• **The probability bound for** $|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)$**:**

By Proposition 2 and assumption A.3, we can show that the expected number of i.i.d. samples belong to U_k for $k \in \mathcal{K}$ is lower bounded with high probability by

$$\begin{aligned} \mathbb{E}_{\boldsymbol{X}}\left[\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k)\right] &\geq p^* C_0(1+\log(t+1)-\log(t_0+1)) \\ &> \frac{1}{2}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right). \end{aligned} \tag{EC.47}$$

Then, we apply the Chernoff inequality (similar to the analysis for (EC.11)) on $\sum_{i=1}^{n} \mathbb{1}(x_i \in U_k)$:

$$\begin{aligned} &\mathbb{P}\left(\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k) < \frac{1}{2}\mathbb{E}_{\boldsymbol{X}}\left[\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k)\right]\right) \leq \exp\left(-\frac{1}{8}\mathbb{E}_{\boldsymbol{X}}\left[\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k)\right]\right) \\ \Rightarrow &\mathbb{P}\left(\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k) < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right) \leq \exp\left(-\frac{1}{16}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right). \end{aligned} \tag{EC.48}$$

When $C_0 \geq 64/p^*$, (EC.48) can be further simplified as follows:

$$\begin{aligned} &\mathbb{P}\left(\sum_{i=1}^{t} \mathbb{1}(\boldsymbol{X}_i \in U_k) < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right) \leq \frac{(t_0+1)^4}{e^4(t+1)^4} \\ \Rightarrow &\mathbb{P}\left(|\mathcal{A}| < \frac{1}{4}p^* C_0 \log\left(\frac{e(t+1)}{t_0+1}\right)\right) \leq \frac{(t_0+1)^4}{e^4(t+1)^4} = \frac{1}{2}\delta_0(t,t_0). \end{aligned} \tag{EC.49}$$

Having proved these two probability bounds, we can combine (EC.45), (EC.46), and (EC.49) to show that

$$\mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_o\right\} \geq 1 - \frac{3}{2}\delta_0(t,t_0),$$

which implies that

$$\begin{aligned} \mathbb{P}\left\{\mathcal{E}_3\right\} &= \mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_s\right\}\mathbb{P}\left\{\mathcal{K}_s\right\} + \mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_o\right\}\mathbb{P}\left\{\mathcal{K}_o\right\} \\ &\geq \mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_o\right\}\mathbb{P}\left\{\mathcal{K}_s\right\} + \mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_o\right\}\mathbb{P}\left\{\mathcal{K}_o\right\} \\ &= \mathbb{P}\left\{\mathcal{E}_3 | k \in \mathcal{K}_o\right\} \geq 1 - \frac{3}{2}\delta_0(t,t_0), \end{aligned} \tag{EC.50}$$

where the first inequality uses the fact that $\mathbb{P}\{\mathcal{E}_3|k\in\mathcal{K}_s\}=1\geq\mathbb{P}\{\mathcal{E}_3|k\in\mathcal{K}_o\}$. Finally, combining (EC.40), (EC.43), and (EC.50), via union bound, we have

$$\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{random}}-\boldsymbol{\beta}^{\text{true}}\|_1\leq\frac{32ns}{|\mathcal{A}|\kappa}\lambda\right)\geq 1-5\delta_0(t,t_0). \tag{EC.51}$$

Moreover, if we pick λ to be small enough (e.g., $\lambda\leq\min\left\{\frac{h\kappa p^*}{3072e\sigma sR_{\max}x_{\max}},\frac{p^*\kappa}{768\sigma sx_{\max}}\right\}$), then when event $\mathcal{E}_3$ holds, we have the following two results:

$$\frac{32ns\lambda}{|\mathcal{A}|\kappa}\leq\frac{32ns\cdot hp^*\kappa}{3072e\sigma sR_{\max}x_{\max}|\mathcal{A}|\kappa}=\frac{h}{4e\sigma R_{\max}x_{\max}}\cdot\frac{n}{|\mathcal{A}|}\cdot\frac{p^*}{24}\leq\frac{h}{4e\sigma R_{\max}x_{\max}} \tag{EC.52}$$

$$\frac{32ns\lambda}{|\mathcal{A}|\kappa}\leq\frac{32nsp^*\kappa}{768\sigma sx_{\max}|\mathcal{A}|\kappa}=\frac{1}{\sigma x_{\max}}\cdot\frac{n}{|\mathcal{A}|}\cdot\frac{p^*}{24}\leq\frac{1}{\sigma x_{\max}}, \tag{EC.53}$$

from which the proposition follows immediately.

Proof of Proposition 5 Here, we will continue using the same requirement of C_0 stated in (EC.35). Because $\{M_k(i)\}$ for $k\in\mathcal{K}$ is a martingale with a bounded difference of 1 per the definition in (19), we can use $M_k(0)$ to bound the value of $M_k(t)$ via Azuma's inequality as follows:

$$\begin{aligned}&\mathbb{P}\left(M_k(t)-M_k(0)\leq-\frac{1}{2}M_k(0)\right)\leq\exp\left(-\frac{M_k(0)^2}{8(t+1)}\right)\\ \Rightarrow&\mathbb{P}\left(M_k(t)\leq\frac{1}{2}M_k(0)\right)\leq\exp\left(-\frac{M_k(0)^2}{8(t+1)}\right).\end{aligned} \tag{EC.54}$$

The $M_k(0)$ term can be stated as follows

$$\begin{aligned}M_k(0)&=\mathbb{E}_{\boldsymbol{X}}\left[\sum_{i=1}^{t}\mathbb{1}(\boldsymbol{X}_i\in U_k,\mathcal{E}_2,\boldsymbol{X}_i\notin\mathcal{R}_{\boldsymbol{x},k})\right]\\&=\sum_{i=1}^{t}\mathbb{P}(\boldsymbol{X}_i\in U_k,\mathcal{E}_2,\boldsymbol{X}_i\notin\mathcal{R}_{\boldsymbol{x},k}).\end{aligned} \tag{EC.55}$$

As $\{\boldsymbol{X}_i\in U_k\}$ is independent of $\{\mathcal{E}_2,\boldsymbol{X}_i\notin\mathcal{R}_{\boldsymbol{x},k}\}$, and $\{\boldsymbol{X}_i\notin\mathcal{R}_{\boldsymbol{x},k}\}$ is independent of $\{\mathcal{E}_2\}$, (EC.55) implies the following inequality

$$\begin{aligned}M_k(0)&=\sum_{i=1}^{t}\mathbb{P}(\boldsymbol{X}_i\in U_k)\mathbb{P}(\mathcal{E}_2)\mathbb{P}(\boldsymbol{X}_i\notin\mathcal{R}_{\boldsymbol{x},k})\\&\geq\sum_{i=1}^{t}p^*\left(1-5\delta_0(t,t_0)\right)\left(1-\frac{2C_0}{t}\right),\end{aligned} \tag{EC.56}$$

where (EC.56) uses assumption A.3, Proposition 4, and the definition of ϵ-decay random sampling scheme with $t_0=2C_0|\mathcal{K}|$.

When $t\geq t_0$, we have

$$5\delta_0(t,t_0)=\frac{10(t_0+1)^4}{e^4(t+1)^4}\leq\frac{1}{2}\text{ and }\frac{2C_0}{t}\leq\frac{1}{2}, \tag{EC.57}$$

where the second inequality uses $t\geq t_0=2C_0|\mathcal{K}|\geq 4C_0$. Inequalities in (EC.57) imply that

$$M_k(0)\geq\sum_{i=1}^{t}\frac{p^*}{4}=\frac{p^*t}{4}. \tag{EC.58}$$

Finally, combining (EC.54) and (EC.58), we can show that the following inequalities hold:

$$\begin{aligned} &\mathbb{P}\left(M_k(t) \le \frac{p^* t}{8}\right) \le \exp\left(-\frac{(p^* t)^2/16}{8(t+1)}\right) \\ \Rightarrow &\mathbb{P}\left(M_k(t) \le \frac{p^* t}{8}\right) \le \exp\left(-\frac{(p^*)^2 t/16}{16}\right) \\ \Rightarrow &\mathbb{P}\left(M_k(t) \le \frac{p^* t}{8}\right) \le \exp\left(-\frac{(p^*)^2 t}{256}\right), \end{aligned} \tag{EC.59}$$

where the second inequality uses $t/(t+1) \ge \frac{1}{2}$.

Proof of Proposition 6 For clear expositions, we first state the following constants:

$$T_0 \ge \max\left\{\frac{48}{C_1 p^*}\log\left(\frac{16}{C_1 p^*}\right), \frac{8}{p^* C_1}\log d, 2|\mathcal{K}|C_0\right\}, \tag{EC.60}$$

$$T_1 \ge \max\left\{\frac{6(192s+\kappa p^* a)^2\lambda_{2,0}^2}{(\kappa p^* \min_{k,i\in\mathcal{S}^k}|\beta_{k,i}^{\text{true}}|)^2}\log\left(\frac{2(192s+\kappa p^* a)^2\lambda_{2,0}^2}{(\kappa p^* \min_{k,i\in\mathcal{S}^k}|\beta_{k,i}^{\text{true}}|)^2}\right), \frac{2(192s+\kappa p^* a)^2\lambda_{2,0}^2\log d}{(\kappa p^* \min_{k,i\in\mathcal{S}^k}|\beta_{k,i}^{\text{true}}|)^2}\right\}, \tag{EC.61}$$

$$\lambda_{2,0} = \frac{4\sigma x_{\max} p^* \kappa a}{p^* \kappa a - 288s}, \tag{EC.62}$$

where C_0, C_1 are defined in (EC.35) and (EC.125) respectively. We further require $a > \frac{1152s}{p^* k} = \mathcal{O}(s)$ in the statement of this proposition, and then we can verify $T_0 = \tilde{\mathcal{O}}(s^2 \log d)$, $T_1 = \tilde{\mathcal{O}}(\beta_{\min}^{-2} s^2 \log d)$, and $\lambda_{2,0} = \mathcal{O}(1)$. Note that if the estimator $\boldsymbol{\beta}_j^{\text{random}}$ is close to $\boldsymbol{\beta}_j^{\text{true}}$ for all $j \in \mathcal{K}$, then assumption A.3 implies that for $\boldsymbol{x}_t \in U_j$, we can clearly separate $\mathbb{E}_\epsilon[R_t|\boldsymbol{x}_t \boldsymbol{\beta}_j^{\text{random}}]$ and $\max_{i\neq j}\mathbb{E}_\epsilon[R_t|\boldsymbol{x}_t \boldsymbol{\beta}_i^{\text{random}}]$. Specifically, part 2 of Lemma EC.3 shows that under event $\mathcal{E}_2$, the following inequality holds for any $\boldsymbol{x} \in U_k$ and $k \in \mathcal{K}_o$:

$$\mathbb{E}_\epsilon[R_k|\boldsymbol{x}^\top \boldsymbol{\beta}_k^{\text{random}}] > \max_{j\neq k}\mathbb{E}_\epsilon[R_j|\boldsymbol{x}^\top \boldsymbol{\beta}_j^{\text{random}}] + \frac{h}{2}, \tag{EC.63}$$

which implies

$$\mathbb{E}_\epsilon[R_k|\boldsymbol{x}_t^\top \boldsymbol{\beta}_k^{\text{random}}] = \max_{j\in\mathcal{K}}\mathbb{E}_\epsilon[R_j|\boldsymbol{x}_t^\top \boldsymbol{\beta}_j^{\text{random}}]$$

and for any $j \neq k$,

$$\mathbb{E}_\epsilon[R_j|\boldsymbol{x}_t^\top \boldsymbol{\beta}_j^{\text{random}}] < \mathbb{E}_\epsilon[R_k|\boldsymbol{x}_t^\top \boldsymbol{\beta}_k^{\text{random}}] - \frac{1}{2}h.$$

Further note that the G-MCP-Bandit algorithm constructs the optimal decision set as follows:

$$\Pi_t = \left\{i : \mathbb{E}_\epsilon[R_i|\boldsymbol{x}_t^\top \boldsymbol{\beta}_i^{\text{random}}] \ge \max_{j\in\mathcal{K}}\mathbb{E}_\epsilon[R_i|\boldsymbol{x}_t^\top \boldsymbol{\beta}_j^{\text{random}}] - \frac{1}{2}h\right\},$$

and therefore, for $\boldsymbol{x}_t \in U_k$, the optimal decision set will be a singleton, i.e., $\Pi_t = \{k\}$, which suggests that decision-makers will assign k as the final decision by merely using the random-sample based estimator $\boldsymbol{\beta}^{\text{random}}$. As the event $\mathcal{E}_2$ is associated with the random estimator using randomly collected samples up to $t-1$ period, the set $\{\boldsymbol{x}_t : \boldsymbol{x}_t \in U_k, \mathcal{E}_2, \boldsymbol{x}_t \notin \mathcal{R}_{\boldsymbol{x},k}\}$ can be viewed as i.i.d. sample from the condition distribution $\mathcal{P}_{\boldsymbol{X}|\boldsymbol{X}\in U_k}$. Then, from Proposition 5, we have

$$\mathbb{P}\left(M_k(t) \le \frac{p^* t}{8}\right) \le \exp\left(-\frac{(p^*)^2 t}{256}\right), \tag{EC.64}$$

where $\{M_k(i)\}$ is defined in (19). As $M_k(t) = \mathbb{E}_{\epsilon,\boldsymbol{X}}\left[\sum_{i=1}^t \mathbb{1}(\boldsymbol{X}_i \in U_k, \mathcal{E}_2, \boldsymbol{X}_i \notin \mathcal{R}_{\boldsymbol{x},k})|\mathcal{F}_t\right] = \sum_{i=1}^t \mathbb{1}(\boldsymbol{x}_i \in U_k, \mathcal{E}_2, \boldsymbol{x}_i \notin \mathcal{R}_{\boldsymbol{x},k})$, the amount of i.i.d. samples in U_k among the whole sample set for arm k up to time t will be lower bounded by $M_k(t)$. Denote $\mathcal{A}$ and n as the set of i.i.d. samples belonging to U_k in the whole sample set and the size of the whole sample, respectively. The following two inequalities hold:

$$\mathbb{P}\left(|\mathcal{A}| \geq \frac{p^* t}{8}\right) \geq 1 - \exp\left(-\frac{(p^*)^2 t}{256}\right) \text{ and } n \leq t. \tag{EC.65}$$

If $|\mathcal{A}| \geq \frac{p^* t}{8}$ and $n \leq t$, then we can obtain the following result:

$$a > \frac{1152 s}{p^* \kappa} \geq \frac{144 st}{|\mathcal{A}|\kappa} > \frac{48 sn}{\kappa|\mathcal{A}|}. \tag{EC.66}$$

Moreover, as $t > T_0 \geq 8(p^* C_1)^{-1} \log d$, then, by (EC.65), we have $|\mathcal{A}| \geq C_1^{-1} \log d$ with high probability. Combining this result with (EC.66) (i.e., two conditions required in Proposition 3), we have the following result via Proposition 3:

$$\begin{aligned}
&\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{whole}} - \boldsymbol{\beta}^{\text{true}}\|_1 \geq \frac{16 ns\zeta}{|\mathcal{A}|\kappa} + \frac{16 ns\rho^{\text{whole}}_{\mathcal{S}/\mathcal{S}_1}}{|\mathcal{A}|\kappa}\lambda\right) \leq \delta_1(n, |\mathcal{A}|, \zeta) + \delta_2(n, |\mathcal{A}|, \lambda) \\
\Rightarrow &\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{whole}} - \boldsymbol{\beta}^{\text{true}}\|_1 \geq \frac{128 s\zeta}{p^*\kappa} + \frac{128 s\rho^{\text{whole}}_{\mathcal{S}/\mathcal{S}_1}}{p^*\kappa}\lambda\right) \leq \delta_1\left(t, \frac{p^* t}{8}, \zeta\right) + \delta_2\left(t, \frac{p^* t}{8}, \lambda\right),
\end{aligned} \tag{EC.67}$$

where (EC.67) uses (EC.65) and the fact that $n \leq t$ in the left-hand side and the facts that $\delta_1(\cdot)$ and $\delta_2(\cdot)$ are monotonically decreasing in $|\mathcal{A}|$ in the right-hand side.

When $t \geq T_1$, (EC.67) can be further simplified. We use Lemma EC.4 in E-Companion with $\alpha = \frac{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2}{2(192 s + \kappa p^* a)^2 \lambda_{2,0}^2}$. When $t \geq T_1$, we have $t \geq 3\alpha^{-1} \log \alpha^{-1}$, combining with the nonnegativity of t, we can show that

$$\begin{aligned}
&\alpha t \geq \log t \\
\Rightarrow &\frac{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2}{2(192 s + \kappa p^* a)^2 \lambda_{2,0}^2} t \geq \log t \\
\Rightarrow &\frac{t}{2} \geq \frac{(192 s + \kappa p^* a)^2 \lambda_{2,0}^2}{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2} \log t.
\end{aligned} \tag{EC.68}$$

Moreover, as $T_1 \geq \frac{2(192 s + \kappa p^* a)^2 \lambda_{2,0}^2 \log d}{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2}$, we can show that when $t > T_1$, the following inequality holds

$$\frac{t}{2} \geq \frac{(192 s + \kappa p^* a)^2 \lambda_{2,0}^2 \log d}{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2}. \tag{EC.69}$$

Combining (EC.69) and (EC.68), we can verify that

$$\begin{aligned}
&t \geq \frac{(192 s + \kappa p^* a)^2 \lambda_{2,0}^2 (\log t + \log d)}{(\kappa p^* \min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}|)^2} \\
\Rightarrow &\min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}| \geq \frac{192 s + \kappa p^* a}{\kappa p^*} \cdot \lambda_{2,0} \sqrt{\frac{\log t + \log d}{t}} \\
\Rightarrow &\min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}| \geq \left(\frac{192 s}{\kappa p^*} + a\right) \cdot \lambda_{2,0} \sqrt{\frac{\log t + \log d}{t}} \\
\Rightarrow &\min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}| \geq \left(\frac{24 ns}{\kappa |\mathcal{A}|} + a\right) \cdot \lambda_{2,0} \sqrt{\frac{\log t + \log d}{t}} \qquad \text{(EC.70)} \\
\Rightarrow &\min_{i:|\beta_i^{\text{true}}|>0} |\beta_i^{\text{true}}| \geq \left(\frac{24 ns}{\kappa |\mathcal{A}|} + a\right) \cdot \lambda_{2,t}. \qquad \text{(EC.71)}
\end{aligned}$$

where (EC.70) uses $|\mathcal{A}| \geq \frac{p^* t}{8}$ and $t \geq n$ and (EC.71) uses the definition of $\lambda_{2,t} = \lambda_{2,0} \cdot \sqrt{(\log t + \log d)/t}$. Now, we use the set $\mathcal{S}_1$ in the Proposition 3 by setting $\lambda = \lambda_{2,t}$, which directly implies that $\mathcal{S}_1 = \mathcal{S}$ so that we have $\mathcal{S}/\mathcal{S}_1$ being the empty set and

$$\rho^{\text{whole}}_{\mathcal{S}/\mathcal{S}_1} = 0. \tag{EC.72}$$

Hence, when $t > T_1$, we can use (EC.72) to simplify (EC.67) into

$$\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{whole}} - \boldsymbol{\beta}^{\text{true}}\|_1 \geq \frac{128 s\zeta}{p^*\kappa}\right) \leq \delta_1\left(t, \frac{p^* t}{8}, \zeta\right) + \delta_2\left(t, \frac{p^* t}{8}, \lambda_{2,t}\right). \tag{EC.73}$$

Finally, we will show that when $t > T_0$, the following two inequalities hold:

$$\delta_1\left(t, \frac{p^* t}{8}, \zeta\right) \leq \frac{2}{(t+1)^2} + 2s\exp\left(-\frac{t\zeta^2}{2\sigma^2 x_{\max}^2}\right), \tag{EC.74}$$

$$\delta_2\left(t, \frac{p^* t}{8}, \lambda_{2,t}\right) \leq \frac{8}{(t+1)^2}. \tag{EC.75}$$

Let's first establish the first inequality (EC.74). Via Lemma EC.4 in E-Companion with $\alpha = \frac{C_1 p^*}{16}$, we can show that because $t > T_0 \geq \max\left\{\frac{48}{C_1 p^*}\log\left(\frac{16}{C_1 p^*}\right), 0\right\}$, we have $\frac{C_1 p^*}{16} t \geq \log t \Rightarrow \frac{C_1 p^*}{8} t \geq 2\log t$, which implies that

$$\exp\left(-C_1|\mathcal{A}|\right) \leq \exp\left(-C_1 \frac{p^*}{8} t\right) \leq \frac{1}{t^2} \leq \frac{2}{(t+1)^2}, \tag{EC.76}$$

where the last inequality uses the fact that $t^{-2} \leq 2(t+1)^{-2}$ holds for all $t \geq T_0 \geq 2|\mathcal{K}|C_0 > 3$. Combining (EC.76) with the definition of $\delta_1(t, |\mathcal{A}|, \zeta)$, we will reach (EC.74). Next, we will show the second inequality (EC.75). When $\lambda_{2,0} = \frac{4\sigma x_{\max} p^* \kappa a}{p^*\kappa a - 288 s}$, we can show that

$$\begin{aligned}
\delta_2(t, \frac{p^* t}{8}, \lambda_{2,t}) &= 4d\exp\left(-\frac{t\lambda_{2,0}^2}{2\sigma^2 x_{\max}^2} \cdot \frac{\log t + \log d}{t} \cdot \left(\frac{1}{2} - \frac{144 s}{p^*\kappa a}\right)^2\right) \\
&= 4d\exp\left(-\frac{16\sigma^2 x_{\max}^2 (p^*\kappa a)^2}{(p^*\kappa a - 288 s)^2} \cdot \frac{t}{2\sigma^2 x_{\max}^2} \cdot \frac{\log t + \log d}{t} \cdot \left(\frac{1}{2} - \frac{144 s}{p^*\kappa a}\right)^2\right) \\
&= 4\exp\left(-2(\log t + \log d) + \log d\right) \leq 4\exp\left(-2\log t\right) \leq \frac{4}{t^2} \leq \frac{8}{(t+1)^2},
\end{aligned}$$

where the last inequality still uses the fact that $t^{-2} \leq 2(t+1)^{-2}$ for $t > 3$. Combining (EC.74) and (EC.75), we have

$$\delta_1\left(t, \frac{p^* t}{8}, \zeta\right) + \delta_2\left(t, \frac{p^* t}{8}, \lambda\right) \leq \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\zeta^2}{2\sigma^2 x_{\max}^2}\right). \tag{EC.77}$$

Proposition 6 directly follows by combining (EC.67), (EC.73), (EC.77), and $\mathbb{P}(\mathcal{E}_2^c) \leq 5\delta_0(T, t_0)$ from Proposition 4.

Proof of Theorem 1 We divide the time, up to time T, into three groups and derive the cumulative regret bound for each group separately. Consider the following three groups:

1. $t \in \{\{t : (\boldsymbol{X}_t, R_t) \in \mathcal{R}_k, k \in \mathcal{K}\} \cup \{t \leq T_0\}\}$.
2. $t \in \{\{t : (\boldsymbol{X}_t, R_t) \notin \mathcal{R}_k, k \in \mathcal{K}, t > T_0\} \cap \{\mathcal{E}_2 \text{ doesn't hold}\}\}$.
3. $t \in \{\{t : (\boldsymbol{X}_t, R_t) \notin \mathcal{R}_k, k \in \mathcal{K}, t > T_0\} \cap \{\mathcal{E}_2 \text{ holds}\}\}$.

In this proof, we follow the same choices of $C_0 = \mathcal{O}(s^2 \log d)$ in (EC.35), $T_1 = \tilde{\mathcal{O}}(\beta_{\min}^{-2} s^2 \log d)$ in (EC.61), $\lambda_1 = \mathcal{O}(s^{-1})$ in (EC.36), $\lambda_{2,0} = \mathcal{O}(1)$ in (EC.62), $T_0 = \tilde{\mathcal{O}}(s^2 \log d)$, and $a > \frac{1152s}{p^*\kappa} = \mathcal{O}(s)$. Beside the requirements for T_0 in (EC.60), we also require that

$$T_0 \geq \max\Big\{t_0, \Big(\frac{512s\lambda_{2,0}}{p^*\kappa}\Big)^2 \log d, 3\Big(\frac{512s\lambda_{2,0}}{p^*\kappa}\Big)^2 \log\Big(\frac{512s\lambda_{2,0}}{p^*\kappa}\Big)^2, \Big(\frac{1024\sigma e^{\sigma x_{\max}} x_{\max} s\lambda_{2,0}}{p^*\kappa}\Big)^2 \log d,$$
$$3\Big(\frac{1024\sigma e^{\sigma x_{\max}} x_{\max} s\lambda_{2,0}}{p^*\kappa}\Big)^2 \log\Big(\frac{1024\sigma e^{\sigma x_{\max}} x_{\max} s\lambda_{2,0}}{p^*\kappa}\Big)^2\Big\}, \tag{EC.78}$$

and T_0 remains on the order of $\tilde{\mathcal{O}}(s^2 \log d)$.

- **Regret in part 1:**

Denote the regret for the first part as $R_1(T)$, and we have

$$R_1(T) \leq R_{\max}\left(\sum_{t=T_0}^{T} \mathbb{1}\left((\boldsymbol{X}_t, R_t) \in \mathcal{R}_k, k \in \mathcal{K}\right) + T_0\right) \leq R_{\max}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k| + T_0\right), \tag{EC.79}$$

where $|\mathcal{R}_k|$ is the cardinality of $\mathcal{R}_k$. From Proposition 2, when $t_0 = 2C_0|\mathcal{K}|$ and $C_0 \geq 20$, we know that

$$\mathbb{P}\left(|\mathcal{R}_k| \leq 3C_0(1 + \log(T) - \log(t_0))\right) \geq 1 - \delta_0(T, t_0)$$
$$\Rightarrow \mathbb{P}\left(|\mathcal{R}_k| \leq 3C_0 \log(T)\right) \geq 1 - \delta_0(T, t_0), \tag{EC.80}$$

which implies

$$\mathbb{P}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k| > |\mathcal{K}| \cdot 3C_0 \log(T)\right) \leq \sum_{k\in\mathcal{K}} \mathbb{P}\left(|\mathcal{R}_k| > 3C_0 \log(T)\right) \leq |\mathcal{K}|\delta_0(T, t_0). \tag{EC.81}$$

We then combine (EC.79) and (EC.81) to bound the regret in part 1:

$$\begin{aligned} R_1(T) \leq R_{\max}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k| + T_0\right) &\leq R_{\max}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k|\right) \mathbb{P}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k| > 3C_0|\mathcal{K}| \log(T)\right) \\ &+ R_{\max}\left(3C_0|\mathcal{K}| \log(T)\right) \mathbb{P}\left(\sum_{k\in\mathcal{K}} |\mathcal{R}_k| \leq 3C_0|\mathcal{K}| \log(T)\right) \\ &+ R_{\max} T_0 \\ &\leq R_{\max} T |\mathcal{K}| \delta_0(T, t_0) + R_{\max} 3C_0 |\mathcal{K}| \log(T) + R_{\max} T_0 \\ &\leq 2R_{\max}|\mathcal{K}|(t_0 + 1) + 3R_{\max} C_0 |\mathcal{K}| \log T + R_{\max} T_0 \qquad \text{(EC.82)} \\ &\leq 3C_0 R_{\max} |\mathcal{K}| \log T + 5R_{\max}|\mathcal{K}| T_0, \qquad \text{(EC.83)} \end{aligned}$$

where (EC.82) uses $T\delta_0(T, t_0) \leq (T+1)\frac{2(t_0+1)^4}{e^4(T+1)^4} \leq 2(t_0 + 1)$ for $T \geq t_0$ and (EC.83) uses $(t_0 + 1) < 2t_0 \leq 2T_0$ and $|\mathcal{K}| > 1$.

- **Regret in part 2:**

Denote the cumulative regret for the second part as $R_2(T)$. From Proposition 4, at time t, we know that

$$\mathbb{P}\left(\|\boldsymbol{\beta}_k^{\text{random}} - \boldsymbol{\beta}_k^{\text{true}}\|_1 \leq \min\left\{\frac{1}{\sigma x_{\max}}, \frac{h}{4e\sigma R_{\max} x_{\max}}\right\}\right) \geq 1 - 5\delta_0(t, t_0),\ k \in \mathcal{K}$$
$$\Rightarrow \mathbb{P}(\mathcal{E}_2(t)^{\text{c}}) \leq 5|\mathcal{K}|\delta_0(t, t_0), \tag{EC.84}$$

where $\mathcal{E}_2(t)$ denotes the event $\mathcal{E}_2$ at time t.

Therefore, $R_2(T)$ can be bounded as follows:

$$
\begin{aligned}
R_2(T) &\leq \mathbb{E}_{\boldsymbol{X},\epsilon}\left[\sum_{i=T_0+1}^{T} \mathbb{1}(\mathcal{E}_2(i)^{\mathrm{c}})R_{\max}\right] = R_{\max}\sum_{i=T_0+1}^{T}\mathbb{P}(\mathcal{E}_2(i)^{\mathrm{c}}) \\
\Rightarrow R_2(T) &\leq R_{\max}\sum_{i=T_0+1}^{T} 5|\mathcal{K}|\delta_0(i,t_0) \\
&\leq 5R_{\max}|\mathcal{K}|\int_{i=T_0}^{T-1}\delta_0(i,t_0)di \\
&= 10R_{\max}|\mathcal{K}|\int_{i=T_0}^{T-1}\frac{(t_0+1)^4}{e^4(i+1)^4}di \\
&= -\frac{10}{3}R_{\max}|\mathcal{K}|\cdot\left.\frac{(t_0+1)^4}{e^4(i+1)^3}\right|_{T_0}^{T-1} \\
&= \frac{10}{3}e^{-4}R_{\max}|\mathcal{K}|(t_0+1)^4(T_0+1)^{-3} - \frac{10}{3}e^{-4}R_{\max}|\mathcal{K}|(t_0+1)^4(T)^{-3} \\
&\leq 2R_{\max}|\mathcal{K}|T_0,
\end{aligned}
$$

where last inequality we use $\frac{10}{3}e^{-4} < \frac{1}{8}$ and $(t_0+1) < 2t_0 \leq 2T_0$.

• **Regret in part 3:**

Denote the cumulative regret for the third part as $R_3(T)$. We first consider the case where $T \leq T_1$. By the second part of Lemma EC.3, it is direct to show that the optimal decision set Π_t constructed in the G-MCP-Bandit Algorithm only contains arms in the optimal decision subset $\mathcal{K}_o$. Without loss of generality, we assume that arm i is the true optimal arm at time t. Then, the regret at time t can be bounded as follows

$$
\begin{aligned}
\mathrm{regret}_t &\leq \mathbb{E}_{\boldsymbol{X}}\left[\sum_{j\in\Pi_t}\mathbb{1}\left(j = \arg\max_{k\in\Pi_t}\mathbb{E}_\epsilon[R_k|\boldsymbol{X}_t^\top\boldsymbol{\beta}_k^{\mathrm{whole}}]\right)\left(\mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{true}}] - \mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{true}}]\right)\right] \\
&\leq \mathbb{E}_{\boldsymbol{X}}\left(\sum_{j\neq i}\mathbb{1}\left(\mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{whole}}] \geq \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{whole}}]\right)\left(\mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{true}}] - \mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{true}}]\right)\right).
\end{aligned}
\tag{EC.85}
$$

We then denote

$$
\mathcal{E}(t,w,\delta_t)_{4,k} = \left\{\mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{true}}] - \mathbb{E}_\epsilon[R_k|\boldsymbol{X}_t^\top\boldsymbol{\beta}_k^{\mathrm{true}}] \in [w\delta_t,(w+1)\delta_t)\right\}, \tag{EC.86}
$$

where $k \neq i, k \in \mathcal{K}_o, w = 0,1,...,$ and $\delta_t > 0$. Then, we have the following bound:

$$
\begin{aligned}
\mathrm{regret}_t &\leq \mathbb{E}_{\boldsymbol{X}}\left(\sum_{w=0}^{w_{1,t}}\sum_{j\neq i}\mathbb{1}\left(\left\{\mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{whole}}] \geq \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{whole}}]\right\}\cap\mathcal{E}(t,w,\delta_t)_{4,j}\right)(w+1)\delta_t\right) \\
&= \sum_{w=0}^{w_{1,t}}(w+1)\delta_t\overbrace{\sum_{j\neq i}\mathbb{P}\left(\left\{\mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{whole}}] \geq \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{whole}}]\right\}\cap\mathcal{E}(t,w,\delta_t)_{4,j}\right)}^{(*)},
\end{aligned}
\tag{EC.87}
$$

where $w_{1,t} = \lceil R_{\max}/\delta_t \rceil$. Now we consider the (*) term in (EC.87), which can be bounded as follows:

$$
\begin{aligned}
(*) &\leq \sum_{j\neq i}\mathbb{P}\left(\left\{\mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{whole}}] - \mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{true}}] \geq \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{whole}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{true}}] + w\delta_t\right\}\cap\mathcal{E}(t,w,\delta_t)_{4,j}\right) \\
&\leq \sum_{j\neq i}\mathbb{P}\left(\left\{\left|\mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{whole}}] - \mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top\boldsymbol{\beta}_j^{\mathrm{true}}]\right| + |\mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{whole}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top\boldsymbol{\beta}_i^{\mathrm{true}}]| \geq w\delta_t\right\}\cap\mathcal{E}(t,w,\delta_t)_{4,j}\right),
\end{aligned}
\tag{EC.88}
$$

where the first inequality uses the fact that $\mathbb{E}_\epsilon[R_i|\boldsymbol{X}_t^\top \boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_j|\boldsymbol{X}_t^\top \boldsymbol{\beta}_j^{\text{true}}] \in [w\delta_t, (w+1)\delta_t)$ when event $\mathcal{E}(t,w,\delta_t)_{4,j}$ holds. To simplify the notation, we denote $\Delta_k = \boldsymbol{\beta}_k^{\text{whole}} - \boldsymbol{\beta}_k^{\text{true}}$ for $k \in \Pi_t$. Combining (EC.88) with the first part of Lemma EC.3, we can show

$$\begin{aligned}(*) &\le \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right\} \cap \mathcal{E}(t,w,\delta_t)_{4,j}\right) \\ &= \sum_{j\neq i} \mathbb{P}\left(\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}), \end{aligned} \tag{EC.89}$$

where the last equality uses the fact that in Δ_i and Δ_j, the terms $\boldsymbol{\beta}_t^{\text{whole}}$ only depend on historical samples upto $t-1$ (independent on t step's information), which implies their independence on $\mathcal{E}(t,w,\delta_t)_{4,j}$.

Denote event $\mathcal{E}_{5,(i,j),t}(w)$ as follows

$$\mathcal{E}_{5,(i,j),t}(w) = \left\{\left\{\|\Delta_i\|_1 \ge \min\left\{\frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}}, 1\right\}\right\} \cup \left\{\|\Delta_j\|_1 \ge \min\left\{\frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}}, 1\right\}\right\}\right\}. \tag{EC.90}$$

Then, conditioning on $\mathcal{E}_{5,(i,j),t}(w)$, the right hand side of (EC.89) can be transformed into

$$\begin{aligned} &\sum_{j\neq i} \mathbb{P}\left(\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &= \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right\} \cap \mathcal{E}_{5,(i,j),t}(w)\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &+ \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right\} \cap (\mathcal{E}_{5,(i,j),t}(w))^{\text{c}}\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &\le \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right\} \cap \mathcal{E}_{5,(i,j),t}(w)\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &+ \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}}\right\} \cap (\mathcal{E}_{5,(i,j),t}(w))^{\text{c}}\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &= \sum_{j\neq i} \mathbb{P}\left(\left\{\|\Delta_j\|_1 + \|\Delta_i\|_1 \ge \frac{w\delta_t}{R_{\max}\sigma e^{\sigma x_{\max}\max\{\|\Delta_j\|_1,\|\Delta_i\|_1\}} x_{\max}}\right\} \cap \mathcal{E}_{5,(i,j),t}(w)\right) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\ &\le \sum_{j\neq i} \mathbb{P}(\mathcal{E}_{5,(i,j),t}(w)) \mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}). \end{aligned} \tag{EC.91}$$

We first bound $\mathbb{P}(\mathcal{E}_{5,(i,j),t}(w))$. As $\mathcal{E}_{5,(i,j),t}(w)$ holds automatically for $w=0$ (i.e., $\mathbb{P}(\mathcal{E}_{5,(i,j),t}(0)) = 1$), we will discuss the remaining cases where $w \ge 1$. As the optimal decision set Π_t only contains arms in the optimal decision subset $\mathcal{K}_o$, from Proposition 6, for $t \ge T_0$, we have the following inequality for $k \in \mathcal{K}_o$:

$$\mathbb{P}\left(\|\Delta_k\|_1 \ge \frac{128 s\zeta}{p^*\kappa} + \frac{128 s \rho^{\text{whole}}_{\mathcal{S}^k/\mathcal{S}^k_{1,t}}}{p^*\kappa}\lambda_{2,t}\right) \le 5\delta_0(t,t_0) + \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\zeta^2}{2\sigma^2 x_{\max}^2}\right). \tag{EC.92}$$

Combining (EC.92) and the choice of T_0 (i.e., $T_0 \ge \max\left\{\left(\frac{512 s\lambda_{2,0}}{p^*\kappa}\right)^2 \log d, 3\left(\frac{512 s\lambda_{2,0}}{p^*\kappa}\right)^2 \log\left(\frac{512 s\lambda_{2,0}}{p^*\kappa}\right)^2\right\}$), we can ensure $\max_i \|\Delta_i\|_1 \le 1$ for all $i \in \mathcal{K}_o$ with high probability for $t > T_0$. To see, note that by setting $\zeta = \lambda_{2,t}$ and using the fact that $\rho^{\text{whole}}_{\mathcal{S}^k/\mathcal{S}^k_{1,t}} \le 1$, we can show that (EC.92) implies

$$\mathbb{P}\left(\|\Delta_k\|_1 \ge \frac{256 s\lambda_{2,t}}{p^*\kappa}\right) \le 5\delta_0(t,t_0) + \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)$$

$$\Rightarrow \mathbb{P}\left(\|\Delta_k\|_1 \geq \frac{256s\lambda_{2,0}}{p^*\kappa}\sqrt{\frac{\log d + \log t}{t}}\right) \leq 5\delta_0(t,t_0) + \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)$$

$$\Rightarrow \mathbb{P}\left(\|\Delta_k\|_1 \geq \frac{256s\lambda_{2,0}}{p^*\kappa}\left(\sqrt{\frac{\log d}{t}} + \sqrt{\frac{\log t}{t}}\right)\right) \leq 5\delta_0(t,t_0) + \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)$$

$$\Rightarrow \mathbb{P}\left(\|\Delta_k\|_1 \geq 1\right) \leq 5\delta_0(t,t_0) + \frac{10}{(t+1)^2} + 2s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right), \tag{EC.93}$$

where (EC.93) uses the facts that $\frac{256s\lambda_{2,0}}{p^*\kappa}\sqrt{\frac{\log d}{t}} \leq \frac{1}{2}$ (because $t > T_0 \geq \left(\frac{512s\lambda_{2,0}}{p^*\kappa}\right)^2 \log d$) and $\frac{256s\lambda_{2,0}}{p^*\kappa}\sqrt{\frac{\log t}{t}} \leq \frac{1}{2}$ (by setting $\alpha = \left(\frac{p^*\kappa}{512s\lambda_{2,0}}\right)^2$ and then using Lemma EC.4 on t for $t > T_0 \geq 3\left(\frac{512s\lambda_{2,0}}{p^*\kappa}\right)^2 \log\left(\frac{512s\lambda_{2,0}}{p^*\kappa}\right)^2$).

We then consider the case with upper bound $\frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}}$ instead of 1 on $\|\Delta_k\|_1$. If we set $\zeta = C_\rho s^{-1} w\delta_t$, where $C_\rho = \frac{p^*\kappa}{256R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}(1+\rho_{\max})}$, $\delta_t = C_{\rho_1} s\sqrt{\frac{\log t + \log d}{t}}$, and $C_{\rho_1} = \frac{512R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}\lambda_{2,0}}{p^*\kappa}$, then we can show that the right-hand-side within the $\mathbb{P}(\cdot)$ term in (EC.92) can be upper bounded as follows:

$$\begin{aligned}
\frac{128s\zeta}{p^*\kappa} + \frac{128s\rho^{\text{whole}}_{\mathcal{S}^k/\mathcal{S}^k_{1,t}}}{p^*\kappa}\lambda_{2,t} &\leq \frac{128s\zeta}{p^*\kappa} + \frac{128s\rho_{\max}}{p^*\kappa}\lambda_{2,t} \\
&= \frac{128sC_\rho s^{-1} w C_{\rho_1} s\sqrt{(\log t + \log d)/t}}{p^*\kappa} + \frac{128s\rho_{\max}}{p^*\kappa}\lambda_{2,t} \\
&= \frac{128sw}{p^*\kappa} \cdot \frac{2}{1+\rho_{\max}} \cdot \lambda_{2,0}\sqrt{\frac{\log t + \log d}{t}} + \frac{128s\rho_{\max}}{p^*\kappa} \cdot \lambda_{2,0}\sqrt{\frac{\log t + \log d}{t}} \\
&= \frac{128}{p^*\kappa}\left(\frac{2w}{1+\rho_{\max}} + \rho_{\max}\right) \cdot \lambda_{2,0} \cdot s\sqrt{\frac{\log t + \log d}{t}} \\
&= \frac{128}{p^*\kappa}\left(\frac{2w}{1+\rho_{\max}} + \rho_{\max}\right) \cdot \lambda_{2,0} \cdot C_{\rho_1}^{-1} \cdot \delta_t \\
&= \left(\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}\right) \cdot \frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}} x_{\max}}
\end{aligned}$$

Note that $\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}$, where $\rho_{\max} \in [0,1]$ and $w \geq 1$, can be upper bounded by 1. To see, we first take the derivative of $\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}$ w.r.t $\rho_{\max}$ to have $-\frac{1}{(1+\rho_{\max})^2} + \frac{1}{2w}$. If $w \geq 2$, then $-\frac{1}{(1+\rho_{\max})^2} + \frac{1}{2w}$ is non-positive for $\rho_{\max} \in [0,1]$, which means that $\rho_{\max} = 0$ is the maximizer for $\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}$, which gives the maximum value of 1; if $1 \leq w < 2$, then $-\frac{1}{(1+\rho_{\max})^2} + \frac{1}{2w}$ will be first negative and then positive for $\rho_{\max} \in [0,1]$, which means that $\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}$ will be maximized at either $\rho_{\max} = 0$ or $\rho_{\max} = 1$, both of which give the maximum value of 1. Therefore, $\frac{1}{1+\rho_{\max}} + \frac{\rho_{\max}}{2w}$ is upper bounded by 1, which implies that

$$\frac{128s\zeta}{p^*\kappa} + \frac{128s\rho^{\text{whole}}_{\mathcal{S}^k/\mathcal{S}^k_{1,t}}}{p^*\kappa}\lambda_{2,t} \leq \frac{w\delta_t}{2R_{\max\,\sigma e^{\sigma x_{\max}} x_{\max}}}. \tag{EC.94}$$

Then, using (EC.93) and (EC.94) for both $\|\Delta_j\|_1$ and $\|\Delta_i\|_1$ for $w \geq 1$, we will have that for $w \geq 1$,

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_{5,(i,j),t}(w)) &\leq \min\left\{1, 10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + \max\left\{4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right), 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right\}\right\} \\
&\leq \min\left\{1, 10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right) + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right\}. \qquad \text{(EC.95)}
\end{aligned}$$

Note that (EC.95) also holds for the case when $w = 0$, as we have $\mathbb{P}(\mathcal{E}_{5,(i,j),t}(0)) = 1$ and $10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right) \geq 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right) = 4s > 1$.

Furthermore, by assumption A.2, we have

$$\mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \le CR_{\max}(1+w)\delta_t. \tag{EC.96}$$

Hence, via (EC.91), (EC.95), and (EC.96), we can show that

$$\begin{aligned}
(*) &\le \sum_{j\neq i} \mathbb{P}(\mathcal{E}_{5,(i,j),t}(w))\mathbb{P}(\mathcal{E}(t,w,\delta_t)_{4,j}) \\
&\le \sum_{j\neq i} \min\left\{1, 10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right) + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right\} \cdot CR_{\max}(1+w)\delta_t.
\end{aligned}$$

Accordingly, the regret bound at time t can be rewritten as follows:

$$\begin{aligned}
\text{regret}_t &\le \sum_{w=0}^{w_{1,t}} (w+1)\delta_t \left(\sum_{j\neq i} \min\left\{1, 10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right)\right.\right. \\
&\qquad\qquad \left.\left. + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right\} \cdot CR_{\max}(1+w)\delta_t\right) \\
&\le CR_{\max}|\mathcal{K}| \sum_{w=0}^{w_{1,t}} (w+1)^2\delta_t^2 \left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right) + \min\left\{1, 4s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right)\right\}\right) \\
&\le CR_{\max}|\mathcal{K}| \left(\overbrace{\left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right) \sum_{w=0}^{w_{1,t}} (1+w)^2\delta_t^2}^{(a)} + \overbrace{\sum_{w=0}^{w_{0,t}} (1+w)^2\delta_t^2}^{(b)} \right. \\
&\qquad\qquad \left. + \overbrace{\sum_{w=w_{0,t}+1}^{w_{1,t}} 4(1+w)^2\delta_t^2 s\exp\left(-\frac{C_\rho^2 t w^2\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\right)}^{(c)} \right),
\end{aligned} \tag{EC.97}$$

where $w_{0,t} = \left\lfloor \sqrt{\frac{2\log(4s)s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}} \right\rfloor$. Next, we will bound part (a), (b) and (c) separately:

$$\begin{aligned}
(a) &< \left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\frac{t\lambda_{2,t}^2}{2\sigma^2 x_{\max}^2}\right)\right)(1+w_{1,t})(1+w_{1,t})^2\delta_t^2 \\
&= \left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-\lambda_{2,0}^2 \frac{\log d + \log t}{2\sigma^2 x_{\max}^2}\right)\right)(1+w_{1,t})(1+w_{1,t})^2\delta_t^2 \\
&= \left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-8\frac{(p^*\kappa a)^2}{(p^*\kappa a - 288s)^2}(\log d + \log t)\right)\right)(1+w_{1,t})(1+w_{1,t})^2\delta_t^2 \\
&= \left(10\delta_0(t,t_0) + \frac{20}{(t+1)^2} + 4s\exp\left(-32(\log d + \log t)\right)\right)(1+w_{1,t})(1+w_{1,t})^2\delta_t^2 \\
&\le \left(10\delta_0(t,t_0) + \frac{24}{(t+1)^2}\right)\delta_t^2(1+w_{1,t})^3 \qquad\qquad \text{(EC.98)} \\
&\le \left(10\delta_0(t,t_0) + \frac{24}{(t+1)^2}\right)\delta_t^2\left(1 + \frac{R_{\max}}{\delta_t} + 1\right)^3 \\
&\le \left(10\delta_0(t,t_0) + \frac{24}{(t+1)^2}\right)(3R_{\max})^3\delta_t^{-1} \qquad\qquad \text{(EC.99)} \\
&\le \frac{540R_{\max}^3(t_0+1)^4}{e^4(t+1)^4\delta_t} + \frac{648R_{\max}^3}{(t+1)^2\delta_t}, \qquad\qquad \text{(EC.100)}
\end{aligned}$$

where (EC.98) uses $s \le d$ and $t^2 \ge t+1$ when $t \ge 1$, and (EC.99) uses $\frac{R_{\max}}{\delta_t} \ge 1$ for $t \ge T_0$, which can be shown in the following analysis

$$
\begin{aligned}
\delta_t &= C_{\rho_1} s \sqrt{\frac{\log d + \log t}{t}} \\
&\le C_{\rho_1} s \sqrt{\frac{\log d + \log T_0}{T_0}} \\
&\le C_{\rho_1} s \sqrt{\frac{\log d}{T_0}} + C_{\rho_1} s \sqrt{\frac{\log T_0}{T_0}} \\
&\le \frac{512 R_{\max} \sigma e^{\sigma x_{\max}} x_{\max} \lambda_{2,0}}{p^* \kappa} s \sqrt{\frac{\log d}{T_0}} + \frac{512 R_{\max} \sigma e^{\sigma x_{\max}} x_{\max} \lambda_{2,0}}{p^* \kappa} s \sqrt{\frac{\log T_0}{T_0}} \\
&\le \frac{R_{\max}}{2} + \frac{512 R_{\max} \sigma e^{\sigma x_{\max}} x_{\max} \lambda_{2,0}}{p^* \kappa} s \sqrt{\frac{\log T_0}{T_0}} \qquad \text{(EC.101)} \\
&\le \frac{R_{\max}}{2} + \frac{512 R_{\max} \sigma e^{\sigma x_{\max}} x_{\max} \lambda_{2,0}}{p^* \kappa} s \left(\frac{p^* \kappa}{1024 \sigma e^{\sigma x_{\max}} x_{\max} s \lambda_{2,0}} \right) \qquad \text{(EC.102)} \\
&\le \frac{1}{2} R_{\max} + \frac{1}{2} R_{\max} = R_{\max},
\end{aligned}
$$

where (EC.101) uses the fact that $T_0 \ge (\frac{1024 \sigma e^{\sigma x_{\max}} x_{\max} s \lambda_{2,0}}{p^* \kappa})^2 \log d$ and (EC.102) uses Lemma EC.4 (by using the fact that $T_0 \ge 3\Big(\frac{1024 \sigma e^{\sigma x_{\max}} x_{\max} s \lambda_{2,0}}{p^* \kappa}\Big)^2 \log\Big(\frac{1024 \sigma e^{\sigma x_{\max}} x_{\max} s \lambda_{2,0}}{p^* \kappa}\Big)^2$ and setting $\alpha = \Big(\frac{p^* \kappa}{1024 \sigma e^{\sigma x_{\max}} x_{\max} s \lambda_{2,0}}\Big)^2$ to show $\alpha T_0 \ge \log T_0$). Next, we can further upper bound part (b) and part (c) in (EC.97) as follows:

$$
(b) < (1 + w_{0,t})(1 + w_{0,t})^2 \delta_t^2 = (1 + w_{0,t})^3 \delta_t^2. \qquad \text{(EC.103)}
$$

$$
\begin{aligned}
(c) &\le 16 s \delta_t^2 \sum_{w = w_{0,t}+1}^{w_{1,t}} w^2 \exp\left(-\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} \cdot w^2 \right) \\
&\le 16 s \delta_t^2 \left(\underbrace{(w_{0,t}+1)^2 \exp\left(-\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} \cdot (w_{0,t}+1)^2 \right)}_{(c_1)} + \underbrace{\sum_{w = w_{0,t}+2}^{w_{1,t}} w^2 \exp\left(-\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} \cdot w^2 \right)}_{(c_2)} \right). \qquad \text{(EC.104)}
\end{aligned}
$$

Let's consider (c_1) and (c_2), separately. As $w_{0,t} = \left\lfloor \sqrt{\frac{2 \log(4s) s^2 \sigma^2 x_{\max}^2}{C_\rho^2 t \delta_t^2}} \right\rfloor$, we have $\sqrt{\frac{2 \log(4s) s^2 \sigma^2 x_{\max}^2}{C_\rho^2 t \delta_t^2}} \le w_{0,t} + 1$, which implies

$$
\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} \cdot (w_{0,t}+1)^2 \ge \frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} \cdot \left(\sqrt{\frac{2 \log(4s) s^2 \sigma^2 x_{\max}^2}{C_\rho^2 t \delta_t^2}} \right)^2 = \log(4s) > 1. \qquad \text{(EC.105)}
$$

Combining (EC.105) with the fact that the function $x \exp(-x)$ is monotonically decreasing for $x \ge 1$, we can show that

$$
\begin{aligned}
&\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} (w_{0,t}+1)^2 \exp\left(-\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} (w_{0,t}+1)^2 \right) \le \log(4s) \exp(-\log 4s) = \frac{\log(4s)}{4s} \\
\Rightarrow &(c_1) = (w_{0,t}+1)^2 \exp\left(-\frac{C_\rho^2 t \delta_t^2}{2 s^2 \sigma^2 x_{\max}^2} (w_{0,t}+1)^2 \right) \le \frac{s \sigma^2 x_{\max}^2 \log(4s)}{2 C_\rho^2 t \delta_t^2}. \qquad \text{(EC.106)}
\end{aligned}
$$

Similarly, as the function $x^2\exp(-x^2)$ is monotonically decreasing for $x\ge 1$, we can upper bound the (c_2) term as follows:

$$
\begin{aligned}
(c_2) &= \sum_{w=w_{0,t}+2}^{w_{1,t}} w^2 \exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot w^2\right)\\
&\le \int_{w_{0,t}+1}^{\infty} w^2 \exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot w^2\right) dw\\
&= \int_{w=w_{0,t}+1}^{w=\infty} -\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\cdot w\cdot d\left[\exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot w^2\right)\right]\\
&= \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\cdot (w_{0,t}+1)\exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot (w_{0,t}+1)^2\right) + \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\int_{w_{0,t}+1}^{\infty} \exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot w^2\right) dw \qquad \text{(EC.107)}\\
&\le \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\left(\frac{w_{0,t}+1}{4s} + \int_{w_{0,t}+1}^{+\infty} \frac{w}{w_{0,t}+1}\exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot w^2\right) dw\right) \qquad \text{(EC.108)}\\
&= \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\left(\frac{w_{0,t}+1}{4s} + \frac{1}{w_{0,t}+1}\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\exp\left(-\frac{C_\rho^2 t\delta_t^2}{2s^2\sigma^2 x_{\max}^2}\cdot (w_{0,t}+1)^2\right)\right)\\
&\le \frac{s\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\left(\frac{w_{0,t}+1}{4s} + \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\cdot\frac{1}{4s}\right)\\
&= \frac{\sigma^2 x_{\max}^2}{4C_\rho^2 t\delta_t^2}\left(w_{0,t}+1+\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\right), \qquad \text{(EC.109)}
\end{aligned}
$$

where (EC.107) uses the integration by parts and (EC.108) uses (EC.105) and $w\ge w_{0,t}+1\ge 1$. Combining (EC.104), (EC.106), and (EC.109), we have

$$
\begin{aligned}
(c) &\le 16s\delta_t^2\left(\frac{s\log(4s)\sigma^2 x_{\max}^2}{2C_\rho^2 t\delta_t^2} + \frac{\sigma^2 x_{\max}^2}{4C_\rho^2 t\delta_t^2}\left(w_{0,t}+1+\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\right)\right)\\
&= \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t}\left(2s\log(4s)+w_{0,t}+1+\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\right). \qquad \text{(EC.110)}
\end{aligned}
$$

Then, combining (EC.97), (EC.100), and (EC.103), (EC.110) with $\delta_t = C_{\rho_1} s\sqrt{\frac{\log t+\log d}{t}}$, we can show that

$$
\begin{aligned}
\text{regret}_t &\le CR_{\max}|\mathcal{K}|\left(\frac{540R_{\max}^3(t_0+1)^4}{e^4(t+1)^4\delta_t} + \frac{648R_{\max}^3}{(t+1)^2\delta_t} + (1+w_{0,t})^3\delta_t^2 + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t}\left(2s\log(4s)+w_{0,t}+1+\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t\delta_t^2}\right)\right)\\
&= CR_{\max}|\mathcal{K}|\left(\frac{540R_{\max}^3(t_0+1)^4}{e^4(t+1)^4\cdot C_{\rho_1} s\sqrt{\frac{\log t+\log d}{t}}} + \frac{648R_{\max}^3}{(t+1)^2\cdot C_{\rho_1} s\sqrt{\frac{\log t+\log d}{t}}} + \frac{(1+w_{0,t})^3\cdot C_{\rho_1}^2 s^2(\log t+\log d)}{t}\right.\\
&\quad\left. + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t}\left(2s\log(4s)+w_{0,t}+1+\frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t C_{\rho_1}^2 s^2(\frac{\log d+\log t}{t})}\right)\right)\\
&\le CR_{\max}|\mathcal{K}|\left(\frac{540R_{\max}^3(t_0+1)^4}{e^4 C_{\rho_1} s\sqrt{\log d}}\cdot\frac{t^{1/2}}{(t+1)^4} + \frac{648R_{\max}^3}{C_{\rho_1} s\sqrt{\log d}}\cdot\frac{t^{1/2}}{(t+1)^2} + \frac{(1+w_{0,t})^3\cdot C_{\rho_1}^2 s^2(\log t+\log d)}{t}\right.\\
&\quad\left. + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t}\left(2s\log(4s)+w_{0,t}+1+\frac{\sigma^2 x_{\max}^2}{C_\rho^2 C_{\rho_1}^2\log d}\right)\right)\\
&\le CR_{\max}|\mathcal{K}|\left(\frac{540R_{\max}^3(t_0+1)^4}{e^4 C_{\rho_1} s\sqrt{\log d}}\cdot\frac{t^{1/2}}{t^4} + \frac{648R_{\max}^3}{C_{\rho_1} s\sqrt{\log d}}\cdot\frac{t^{1/2}}{t^{3/2}\cdot\sqrt{T_0+1}} + \frac{(1+w_{0,t})^3 C_{\rho_1}^2\cdot s^2(\log t+\log d)}{t}\right.\\
&\quad\left. + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t}\left(2s\log(4s)+w_{0,t}+1+\frac{\sigma^2 x_{\max}^2}{C_\rho^2 C_{\rho_1}^2\log d}\right)\right)
\end{aligned}
$$

$$\leq R_{\max}|\mathcal{K}|\left(\frac{540CR_{\max}^3(t_0+1)^4}{e^4C_{\rho_1}s(\log^{1/2}d)t^{7/2}}+\frac{C_3+CC_{\rho_1}^2(1+w_{0,t})^3s^2\log t}{t}\right), \tag{EC.111}$$

where

$$C_3=\frac{648CR_{\max}^3}{C_{\rho_1}s(\log^{1/2}d)\sqrt{T_0+1}}+CC_{\rho_1}^2(1+w_{0,t})^3s^2\log d+\frac{4Cs\sigma^2x_{\max}^2}{C_\rho^2}\left(2s\log(4s)+w_{0,t}+1+\frac{\sigma^2x_{\max}^2}{C_\rho^2C_{\rho_1}^2\log d}\right). \tag{EC.112}$$

Hence, the third part of the regret can be bounded as follows:

$$\begin{aligned}
R_3(T)&\leq\sum_{t=T_0+1}^{T}\text{regret}_t\cdot\mathbb{P}(\mathcal{E}_2(t))\\
&\leq\sum_{t=T_0+1}^{T}R_{\max}|\mathcal{K}|\left(\frac{540CR_{\max}^3(t_0+1)^4}{e^4C_{\rho_1}s\log^{1/2}dt^{7/2}}+\frac{C_3+CC_{\rho_1}^2(1+\max_{t\leq T}w_{0,t})^3s^2\log t}{t}\right)\\
&\leq\int_{T_0}^{T}R_{\max}|\mathcal{K}|\left(\frac{540CR_{\max}^3(t_0+1)^4}{e^4C_{\rho_1}s\log^{1/2}dt^{7/2}}+\frac{C_3+CC_{\rho_1}^2(1+\max_{t\leq T}w_{0,t})^3s^2\log t}{t}\right)dt\\
&\leq R_{\max}|\mathcal{K}|\left(-\frac{1080CR_{\max}^3(t_0+1)^4}{5e^4C_{\rho_1}s(\log^{1/2}d)t^{5/2}}+C_3\log t+CC_{\rho_1}^2(1+\max_{t\leq T}w_{0,t})^3s^2\log^2t\right)\Bigg|_{T_0}^{T}\\
&\leq R_{\max}|\mathcal{K}|\left(\frac{1080CR_{\max}^3(t_0+1)^4}{5e^4C_{\rho_1}s(\log^{1/2}d)(T_0)^{5/2}}+C_3\log T+CC_{\rho_1}^2(1+\max_{t\leq T}w_{0,t})^3s^2\log^2T\right)\\
&\leq R_{\max}|\mathcal{K}|\left(C_4(t_0+1)+C_3\log T+C_5\log^2T\right)\\
&\leq R_{\max}|\mathcal{K}|\left(2C_4T_0+C_3\log T+C_5\log^2T\right),
\end{aligned} \tag{EC.113}$$

where we set

$$C_4=\frac{216CR_{\max}^3(t_0+1)^3}{e^4C_{\rho_1}s(\log^{1/2}d)(T_0)^{5/2}} \tag{EC.114}$$

$$C_5=CC_{\rho_1}^2(1+\max_{t\leq T}w_{0,t})^3s^2 \tag{EC.115}$$

and $t_0\leq T_0$. As we set $t_0=2C_0|\mathcal{K}|$ and $C_0=\mathcal{O}(s^2\log d)$, which implies $C_4=\mathcal{O}(1)$. Moreover, as $w_{0,t}=\left\lfloor\sqrt{\frac{2s^2\log(4s)\sigma^2x_{\max}^2}{C_\rho^2t\delta_t^2}}\right\rfloor\leq\left\lfloor\sqrt{\frac{\log(4s)\sigma^2x_{\max}^2}{2\lambda_{2,0}^2\log d}}\right\rfloor=\mathcal{O}(1)$ and $C_\rho=\mathcal{O}((1+\rho_{\max})^{-1})$, we can directly show that $C_3\leq\tilde{\mathcal{O}}((1+\rho_{\max})^2s^2\log d)$ and $C_5=\mathcal{O}(s^2)$.

Next, we consider the other case where $T>T_1$. Via proposition 6, we know that when $T>T_1$, for all $k\in\mathcal{K}$, $\mathcal{S}_{1,t}^k=\mathcal{S}^k\Rightarrow\rho_{\mathcal{S}^k/\mathcal{S}_{1,t}^k}=0$. In this case, we can restate (EC.92) as follows

$$\mathbb{P}\left(\|\Delta_k\|_1\geq\frac{128s\zeta}{p^*\kappa}\right)\leq5\delta_0(t,t_0)+\frac{10}{(t+1)^2}+2s\exp\left(-\frac{t\zeta^2}{2\sigma^2x_{\max}^2}\right).$$

Then, by setting $\zeta=C_\rho s^{-1}w\delta_t$, via the similar analysis to (EC.94), we can show

$$\begin{aligned}
\frac{128s\zeta}{p^*\kappa}&=\frac{128s}{p^*\kappa}\cdot\frac{p^*\kappa}{256R_{\max}\sigma e^{\sigma x_{\max}}x_{\max}(1+\rho_{\max})}\cdot s^{-1}w\delta_t\\
&=\frac{1}{1+\rho_{\max}}\cdot\frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}}x_{\max}}\\
&\leq\frac{w\delta_t}{2R_{\max}\sigma e^{\sigma x_{\max}}x_{\max}},
\end{aligned}$$

which implies that (EC.95) still holds and we have the same separation as in (EC.97). The analyses for parts (a), (b), and (c) remain unchanged. As we choose $\delta_t=C_{\rho_1}s\sqrt{\frac{\log d}{t}}$, which is different from the choice in the

$T \le T_1$ case where $\delta_t = C_{\rho_1} s\sqrt{\frac{\log d + \log t}{t}}$, the regret$_t$ calculation will be slightly different from the analysis for (EC.111). In particular, we can show that

$$\begin{aligned}
\text{regret}_t &\le C R_{\max}|\mathcal{K}| \left(\frac{540 R_{\max}^3 (t_0+1)^4}{e^4 (t+1)^4 \delta_t} + \frac{540 R_{\max}^3}{(t+1)^2 \delta_t} + (1+w_{0,t})^3 \delta_t^2 + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t} \left(2s\log(4s) + w_{0,t} + 1 + \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t \delta_t^2} \right) \right) \\
&= C R_{\max}|\mathcal{K}| \left(\frac{540 R_{\max}^3 (t_0+1)^4}{e^4 (t+1)^4 \cdot C_{\rho_1} s\sqrt{\frac{\log d}{t}}} + \frac{540 R_{\max}^3}{(t+1)^2 \cdot C_{\rho_1} s\sqrt{\frac{\log d}{t}}} + \frac{(1+w_{0,t})^3 \cdot C_{\rho_1}^2 s^2 \log d}{t} \right. \\
&\quad \left. + \frac{4s\sigma^2 x_{\max}^2}{C_\rho^2 t} \left(2s\log(4s) + w_{0,t} + 1 + \frac{s^2\sigma^2 x_{\max}^2}{C_\rho^2 t C_{\rho_1}^2 s^2 (\frac{\log d}{t})} \right) \right) \\
&\le R_{\max}|\mathcal{K}| \left(\frac{540 C R_{\max}^3 (t_0+1)^4}{e^4 C_{\rho_1} s (\log^{1/2} d) t^{7/2}} + \underbrace{\frac{C_3 + C C_{\rho_1}^2 (1+w_{0,t})^3 s^2}{t}}_{(d)} \right). \qquad \text{(EC.116)}
\end{aligned}$$

When comparing (EC.111) to (EC.116), we can show that the $\log t$ term disappeared in the (d) term. Therefore, following similar analysis as in (EC.113), the third part of the regret, when $T > T_1$, can be upper bounded as follows:

$$R_3(T) \le R_{\max}|\mathcal{K}|(2C_4 T_0 + (C_3 + C_5)\log T + C_5 \log^2 T_1).$$

Finally, the total regret bound can be obtained by combining the bounds from all three parts: when $T \le T_1$, we have

$$\begin{aligned}
R_1(T) + R_2(T) + R_3(T) &\le 3C_0 R_{\max}|\mathcal{K}|\log T + 5R_{\max}|\mathcal{K}|T_0 + 2R_{\max}|\mathcal{K}|T_0 + R_{\max}|\mathcal{K}|(2C_4 T_0 + C_3 \log T + C_5 \log^2 T) \\
&= R_{\max}|\mathcal{K}| \left[(3C_0 + C_3)\log T + C_5 \log^2 T + (7 + 2C_4) T_0 \right] \\
&= \tilde{\mathcal{O}}(s^2 (\log d + \log T)\log T);
\end{aligned}$$

when $T > T_1$, we have

$$\begin{aligned}
&R_1(T) + R_2(T) + R_3(T) \\
&\le 3C_0 R_{\max}|\mathcal{K}|\log T + 5R_{\max}|\mathcal{K}|T_0 + 2R_{\max}|\mathcal{K}|T_0 + R_{\max}|\mathcal{K}|(2C_4 T_0 + (C_3 + C_5)\log T + C_5 \log^2 T_1) \\
&= R_{\max}|\mathcal{K}| \left[(3C_0 + C_3 + C_5)\log T + (7 + 2C_4) T_0 + C_5 \log^2 T_1 \right] \\
&= \tilde{\mathcal{O}}(s^2 \log d \log T).
\end{aligned}$$

Proof of Theorem 2 We adopt the FISTA method in Beck and Teboulle (2009) as the Lasso solver in the 2sWL procedure. For completeness, we first present the FISTA method in our settings.

FISTA Method:

Require: Loss function $\mathcal{L}(\boldsymbol{\beta})$, penalty parameter λ, total iteration number $k_0 \ge 1$, initial solution $\boldsymbol{\beta}_0$, and step-size l_0.

Step 0: Set $\boldsymbol{y}_1 = \boldsymbol{\beta}_0$, $t_1 = 1$, and $k = 1$.

While $k \le k_0$:

$$\boldsymbol{\beta}_k = \arg\min_{\boldsymbol{\beta}} \left\{ \lambda\|\boldsymbol{\beta}\|_1 + \frac{l_0}{2} \left\| \boldsymbol{\beta} - \left(\boldsymbol{y}_k - \frac{1}{l_0}\nabla\mathcal{L}(\boldsymbol{y}_k) \right) \right\|_2^2 \right\} \qquad (*)$$

$$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$$

$$\boldsymbol{y}_{k+1} = \boldsymbol{\beta}_k + \frac{t_k - 1}{t_{k+1}} \cdot (\boldsymbol{\beta}_k - \boldsymbol{\beta}_{k-1})$$

$$k = k+1$$

The major computation cost in FISTA is from part ($*$), in which we need the full gradient $\nabla\mathcal{L}(\boldsymbol{y_k})$. By definition, we have $\mathcal{L}(\boldsymbol{\beta})=\frac{1}{n}\sum_{j=1}^n f(R_j|\boldsymbol{X}_j^T\boldsymbol{\beta})$. Hence, to evaluate each $\nabla\mathcal{L}(\boldsymbol{y_k})$, we need to compute $\mathcal{O}(dn)$ scalers multiplication. Let $\boldsymbol{\beta}^*$ be the optimal solution, $\boldsymbol{\beta}_0$ be the initial solution, and the Theorem 4.4 in Beck and Teboulle (2009) implies that for any $k\geq 1$

$$\mathcal{L}(\boldsymbol{\beta}_k)+\lambda\|\boldsymbol{\beta}_k\|_1-\mathcal{L}(\boldsymbol{\beta}^*)+\lambda\|\boldsymbol{\beta}^*\|_1\leq\mathcal{O}\left(\frac{L\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^*\|_2^2}{(k+1)^2}\right), \tag{EC.117}$$

where L is the Lipschitz constant of $\mathcal{L}(\boldsymbol{\beta})$ function and will be on the order of $\mathcal{O}(x_{\max}bd)$. Therefore, to achieve ϵ-optimal solution, the required total iterations k_0 will be on the order of $\mathcal{O}(x_{\max}b\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^*\|_2^2\epsilon^{-1/2})=\mathcal{O}(x_{\max}b^3\epsilon^{-1/2})$, where we use $\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^*\|_2^2\leq 4b^2d^2$ by using $\|\boldsymbol{\beta}\|_1\leq b$ for all feasible $\boldsymbol{\beta}$ in assumption A.1. Therefore, the total computational cost for running FISTA becomes $\mathcal{O}(x_{\max}b^3dn\epsilon^{-1/2})$. Note that at step T, each arm can not be pulled more than T times, so the maximum computation cost of FISTA will be $\mathcal{O}(x_{\max}b^3d^4T\epsilon^{-1/2})$.

Next, we will upper bound the total number of the FISTA method called by time T. At each step, the G-MCP-Bandit algorithm will require to update $\boldsymbol{\beta}_k^{\text{random}}$ and $\boldsymbol{\beta}_k^{\text{whole}}$ by 2sWL for $k\in\mathcal{K}$ and each 2sWL procedure will need to run FISTA two times. So the average computation cost will be

$$\text{Average computation cost}\leq\mathcal{O}\left(\frac{1}{T}\cdot\sum_{k\in\mathcal{K}}\sum_{t=1}^{T}2x_{\max}b^3dt\epsilon^{-1/2}\right)=\mathcal{O}(|\mathcal{K}|x_{\max}b^3d^4T\epsilon^{-1/2}). \tag{EC.118}$$

Next, we consider the long-run computation cost. We can reduce the computation cost with a warm start from the previous step. Via Proposition 3, we can show that with high probability, for $T\geq\max\{T_1,t_0^2\}$ and $\zeta=\frac{\epsilon^2p^*\kappa}{16s}$, the following inequality holds:

$$\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{MCP}}-\boldsymbol{\beta}^{\text{true}}\|_1\leq\epsilon^{\frac{1}{4}}\right)\geq 1-\frac{10}{(T+1)^2}-2\exp\left(-\frac{(T+1)(\epsilon^{1/4}p^*\kappa)^2}{512s^2\sigma^2x_{\max}^2}+\log s\right). \tag{EC.119}$$

Moreover, when $T\geq\max\left\{\frac{1024s^2\log(s)\sigma^2x_{\max}^2}{\epsilon^{1/2}(p^*\kappa)^2},\frac{6144s^2\sigma^2x_{\max}^2}{\epsilon^{1/2}(p^*\kappa)^2}\log\left(\frac{2048s^2\sigma^2x_{\max}^2}{\epsilon^{1/2}(p^*\kappa)^2}\right)\right\}=\mathcal{O}(s^2\log(s)\epsilon^{-1/2})$, via Lemma EC.4, we have

$$\begin{aligned}&\frac{(T+1)\epsilon^{1/2}(p^*\kappa)^2}{1024s^2\sigma^2x_{\max}^2}\geq\log(s)\text{ and }\frac{(T+1)\epsilon^{1/2}(p^*\kappa)^2}{1024s^2\sigma^2x_{\max}^2}\geq 2\log(T+1)\\ \Rightarrow &2\exp\left(-\frac{(T+1)(\epsilon^{1/4}p^*\kappa)^2}{512s^2\sigma^2x_{\max}^2}+\log s\right)\leq\frac{1}{(T+1)^2}.\end{aligned} \tag{EC.120}$$

Combining (EC.119) and (EC.120), we have

$$\mathbb{P}\left(\|\boldsymbol{\beta}^{\text{MCP}}-\boldsymbol{\beta}^{\text{true}}\|_1\leq\epsilon^{1/4}\right)\geq 1-\mathcal{O}(T^{-2}) \tag{EC.121}$$

and via Proposition 5 and Lemma EC.2 in E-Companion, we can show that for $T\geq\mathcal{O}(s^2\log d\epsilon^{-1/2})$, similar result holds

$$\mathbb{P}(\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1\leq\epsilon^{1/4})\geq 1-\mathcal{O}(T^{-2}). \tag{EC.122}$$

Note that (EC.121) and (EC.122) imply that for large enough T, both previous step solution and current step solution are close to $\boldsymbol{\beta}^{true}$ with high probability. If we use the previous step solution to initialize the FISTA algorithm, then we have

$$\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^*\|_2^2\leq\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^*\|_1^2=\|\boldsymbol{\beta}_0-\boldsymbol{\beta}^{\text{true}}+\boldsymbol{\beta}^{\text{true}}-\boldsymbol{\beta}^*\|_1^2$$

$$\leq 2\|\boldsymbol{\beta}_0 - \boldsymbol{\beta}^{\text{true}}\|_1^2 + 2\|\boldsymbol{\beta}^* - \boldsymbol{\beta}^{\text{true}}\|_1^2$$
$$\leq 4\epsilon^{1/2}, \tag{EC.123}$$

where the first inequality uses $(a+b)^2 \leq 2a^2 + 2b^2$, and the last inequality is because $\boldsymbol{\beta}_0$ is initialized with previous step's solution and $\boldsymbol{\beta}^*$ is the current step's MCP solution. Therefore, the results in (EC.118) can be improved to $\mathcal{O}(|\mathcal{K}|x_{\max}bd^2T)$.

EC.2. Appendix: Supplemental Lemmas and Proofs

LEMMA EC.1. *Let n be the size of the whole sample set and $\mathcal{A}$ be the random i.i.d. sample set consisting of $\boldsymbol{X} \in \mathbf{R}^d$ for $k \in \mathcal{K}_s$ and $\boldsymbol{X} \in U_k$ for $k \in \mathcal{K}_o$. Under assumptions A.1, A.4, and A.5, when $|\mathcal{A}| \geq C_1^{-1} \log d$, the follow inequality holds for all feasible $\boldsymbol{\xi}$ and $\boldsymbol{u}$ such that $\|\boldsymbol{u}_{\mathcal{S}^c}\|_1 \leq 3\|\boldsymbol{u}_{\mathcal{S}}\|_1$:*

$$\mathbb{P}\left(\frac{|\mathcal{A}|\kappa}{2ns}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \leq \boldsymbol{u}^\top \nabla^2 \mathcal{L}(\boldsymbol{\xi})\boldsymbol{u}\right) \geq 1 - \exp(-C_1|\mathcal{A}|), \tag{EC.124}$$

where

$$C_1 = \min\left\{1, \kappa^2 / \left(192 s \sigma_2 x_{\max}^2 (2 + \sqrt{\sigma_2} x_{\max})\right)^2\right\}. \tag{EC.125}$$

Proof of EC.1 Let $\mathcal{L}_{\mathcal{A}}(\boldsymbol{\beta})$ be the loss function with the sample set $\mathcal{A}$. Denote $\boldsymbol{Z}_j = \boldsymbol{X}_j \sqrt{f''_{yy}(R_j|\boldsymbol{X}_j^\top \boldsymbol{\xi})}$, where we replace r and y in $f''_{yy}(r|y)$ by R_j and $\boldsymbol{X}_j^\top \boldsymbol{\xi}$ respectively. We then can present $\nabla^2 \mathcal{L}_{\mathcal{A}}(\boldsymbol{\xi})$ as follows:

$$\nabla^2 \mathcal{L}_{\mathcal{A}}(\boldsymbol{\xi}) = \frac{1}{|\mathcal{A}|}\sum_{j \in \mathcal{A}} \boldsymbol{X}_j \boldsymbol{X}_j^\top f''_{yy}(R_j|\boldsymbol{X}_j^\top \boldsymbol{\xi}) = \frac{1}{|\mathcal{A}|}\sum_{j \in \mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top.$$

As all realization of $\boldsymbol{X}$ is element-wise bounded by $x_{\max}$ (see assumption A.1) and $f''_{yy}(r_j|\boldsymbol{x}_j^\top \boldsymbol{\xi}) \leq \sigma_2$ (see assumptions A.4), $\boldsymbol{Z}_j$ is element-wise bounded by $\|\boldsymbol{Z}_j\|_\infty = \left\|\boldsymbol{X}_j \sqrt{f''_{yy}(R_j|\boldsymbol{X}_j^\top \boldsymbol{\xi})}\right\|_\infty \leq \sqrt{\sigma_2} x_{\max} \doteq z_{\max}$. Then, we use Bühlmann and Van De Geer (2011) to build the connection between the sample matrix $\frac{1}{|\mathcal{A}|}\sum_{j \in \mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top$ and its population counterpart $\mathbb{E}_{\boldsymbol{Z}}[\boldsymbol{Z}_j \boldsymbol{Z}_j^\top]$. By setting $K = z_{\max}$ and $\sigma_0 = \sqrt{2} z_{\max}$ in the exercise 14.3 in Bühlmann and Van De Geer (2011), for $t > 0$, we have

$$P\left\{\left\|\frac{1}{|\mathcal{A}|}\sum_{j \in \mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top - \mathbb{E}_{\boldsymbol{Z}}[\boldsymbol{Z}_j \boldsymbol{Z}_j^\top]\right\|_\infty \geq 2z_{\max}^2 t + 4z_{\max}^2 \sqrt{t} + \sqrt{8} z_{\max}^2 \lambda\left(\frac{\sqrt{2}}{2}, |\mathcal{A}|, \binom{d}{2}\right)\right\} \leq \exp\left(-|\mathcal{A}|t\right), \tag{EC.126}$$

where $\lambda\left(\frac{\sqrt{2}}{2}, |\mathcal{A}|, \binom{d}{2}\right) = \sqrt{\frac{2\log(d(d-1))}{|\mathcal{A}|}} + \frac{z_{\max}\log(d(d-1))}{|\mathcal{A}|}$.

When $|\mathcal{A}| \geq C_1^{-1} \log d$ and $t = C_1$ in (EC.126), the following inequalities hold:

$$\begin{aligned} 2z_{\max}^2 t + 4z_{\max}^2 \sqrt{t} &= 2z_{\max}^2 C_1 + 4z_{\max}^2 \sqrt{C_1} \\ &\leq 6z_{\max}^2 \sqrt{C_1} \qquad \text{(EC.127)} \\ \sqrt{8} z_{\max}^2 \lambda\left(\frac{\sqrt{2}}{2}, |\mathcal{A}|, \binom{d}{2}\right) &\leq \sqrt{8} z_{\max}^2 \left(\sqrt{\frac{2\log(d^2)}{|\mathcal{A}|}} + \frac{z_{\max}\log(d^2)}{|\mathcal{A}|}\right) \\ &= \sqrt{8} z_{\max}^2 \left(\sqrt{\frac{4\log d}{|\mathcal{A}|}} + \frac{2z_{\max}\log d}{|\mathcal{A}|}\right) \\ &\leq \sqrt{8} z_{\max}^2 \left(2\sqrt{C_1} + 2z_{\max} C_1\right) \\ &\leq 4\sqrt{2} z_{\max}^2 (1 + z_{\max})\sqrt{C_1}, \qquad \text{(EC.128)} \end{aligned}$$

where in (EC.127) and (EC.128) we use the fact that when $C_1 \leq 1$, we have $\sqrt{C_1} \geq C_1$. Combining (EC.127) and (EC.128), we have

$$\begin{aligned} 2z_{\max}^2 t + 4z_{\max}^2 \sqrt{t} + \sqrt{8} z_{\max}^2 \lambda\left(\frac{\sqrt{2}}{2}, |\mathcal{A}|, \binom{d}{2}\right) &\leq 2z_{\max}^2 \left(3 + 2\sqrt{2}(1 + z_{\max})\right)\sqrt{C_1} \\ &< 6z_{\max}^2 \left(2 + z_{\max})\right)\sqrt{C_1} \leq \frac{\kappa}{32s}, \qquad \text{(EC.129)} \end{aligned}$$

where (EC.129) uses $\sqrt{2} \le \frac{3}{2}$ and $C_1 \le \kappa^2 / \left(192 s \sigma_2 x_{\max}^2 \left(2 + \sqrt{\sigma_2} x_{\max}\right)\right)^2$. Combining (EC.126) and (EC.129), we have

$$\mathbb{P}\left\{\left\|\frac{1}{|\mathcal{A}|}\sum_{j\in\mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top - \mathbb{E}_{\boldsymbol{Z}}[\boldsymbol{Z}_j \boldsymbol{Z}_j^\top]\right\|_\infty \le \frac{\kappa}{32s}\right\} \ge 1 - \exp\left(-C_1 |\mathcal{A}|\right). \tag{EC.130}$$

By the definition of $\boldsymbol{Z}_j$, we can verify that $\mathbb{E}_{\boldsymbol{Z}}[\boldsymbol{Z}_j \boldsymbol{Z}_j^\top] = \mathbb{E}_{\boldsymbol{X},\epsilon}[f''_{yy}(R_j|\boldsymbol{\xi}^\top \boldsymbol{X}_j)\boldsymbol{X}_j \boldsymbol{X}_j^\top]$. Via assumption A.5, we have restricted eigenvalue condition holds for $\mathbb{E}_{\boldsymbol{Z}}[\boldsymbol{Z}_j \boldsymbol{Z}_j^\top]$ with parameter κ. Combining (EC.130) with the Corollary 6.8 in Bühlmann and Van De Geer (2011), we set $\tilde{\lambda} = \frac{\kappa}{32s}$ in Corollary 6.8 to show $\frac{1}{|\mathcal{A}|}\sum_{j\in\mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top$ also has restricted eigenvalue condition with parameter $\kappa/2$, which implies

$$\begin{aligned} &\mathbb{P}\left(\frac{\kappa}{2s}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \le \frac{1}{|\mathcal{A}|}\sum_{j\in\mathcal{A}} \boldsymbol{Z}_j \boldsymbol{Z}_j^\top\right) \ge 1 - \exp(-C_1|\mathcal{A}|) \\ \Rightarrow &\mathbb{P}\left(\frac{\kappa}{2s}\|\boldsymbol{u}_{\mathcal{S}}\|_1^2 \le \boldsymbol{u}^\top \nabla^2 \mathcal{L}_{\mathcal{A}}(\boldsymbol{\xi})\boldsymbol{u}\right) \ge 1 - \exp(-C_1|\mathcal{A}|). \end{aligned} \tag{EC.131}$$

Note that for any realizations $\{\boldsymbol{x}_j, r_j\}$ of $\{\boldsymbol{X}_j, R_j\}$, we have

$$\begin{aligned} \boldsymbol{u}^T \nabla^2 \mathcal{L}_{\mathcal{A}}(\boldsymbol{\xi})\boldsymbol{u} &= \boldsymbol{u}^T \left[\frac{1}{n}\sum_{j\in\mathcal{A}} x_j x_j^\top f''_{yy}(r_j|\boldsymbol{x}_j^T \boldsymbol{\xi})\right]\boldsymbol{u} + \boldsymbol{u}^T \left[\frac{1}{n}\sum_{j\in(\mathcal{A})^c} x_j x_j^\top f''_{yy}(r_j|\boldsymbol{x}_j^T \boldsymbol{\xi})\right]\boldsymbol{u} \\ &\ge \boldsymbol{u}^T \left[\frac{1}{n}\sum_{j\in\mathcal{A}} x_j x_j^\top f''_{yy}(r_j|\boldsymbol{x}_j^T \boldsymbol{\xi})\right]\boldsymbol{u} = \frac{|\mathcal{A}|}{n}\boldsymbol{u}^\top \nabla^2 \mathcal{L}_{\mathcal{A}}(\boldsymbol{\xi})\boldsymbol{u}. \end{aligned} \tag{EC.132}$$

The desirable result follows directly by combining (EC.131) and (EC.132).

LEMMA EC.2. *Let n be the size of the whole sample set and $\mathcal{A}$ be the random i.i.d. sample set consisting of $\boldsymbol{X} \in \mathbf{R}^d$ for $k \in \mathcal{K}_s$ and $\boldsymbol{X} \in U_k$ for $k \in \mathcal{K}_o$. Per assumptions A.1, A.4 and A.5, when $|\mathcal{A}| \ge C_1^{-1}\log d$, the follow result holds with probability at least $1 - \exp\left(-C_1|\mathcal{A}|\right) - 2d\exp\left(-\frac{n\lambda^2}{8\sigma^2 x_{\max}^2}\right)$:*

$$\|\boldsymbol{\beta}^{\text{lasso}} - \boldsymbol{\beta}^{\text{true}}\|_1 \le \frac{24 n s \lambda}{|\mathcal{A}|\kappa}, \tag{EC.133}$$

where $\boldsymbol{\beta}^{\text{lasso}}$ is the lasso estimator and C_1 is defined in (EC.125)

Proof of lemma EC.2 We first show that $\|\boldsymbol{\beta}_{\mathcal{S}^c}^{\text{lasso}} - \boldsymbol{\beta}_{\mathcal{S}^c}^{\text{true}}\|_1 \le 3\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}} - \boldsymbol{\beta}_{\mathcal{S}}^{\text{true}}\|_1$ holds. As $\boldsymbol{\beta}^{\text{lasso}}$ is the optimal solution, we have

$$\mathcal{L}(\boldsymbol{\beta}^{\text{lasso}}) + \lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1 \le \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \lambda\|\boldsymbol{\beta}^{\text{true}}\|_1$$

$$\mathcal{L}(\boldsymbol{\beta}^{\text{lasso}}) - \mathcal{L}(\boldsymbol{\beta}^{\text{true}}) + \lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1 \le \lambda\|\boldsymbol{\beta}^{\text{true}}\|_1 \tag{EC.134}$$

$$\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^T(\boldsymbol{\beta}^{\text{lasso}} - \boldsymbol{\beta}^{\text{true}}) + \lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1 \le \lambda\|\boldsymbol{\beta}^{\text{true}}\|_1 \tag{EC.135}$$

$$-\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty \|\boldsymbol{\beta}^{\text{lasso}} - \boldsymbol{\beta}^{\text{true}}\|_1 + \lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1 \le \lambda\|\boldsymbol{\beta}^{\text{true}}\|_1, \tag{EC.136}$$

where (EC.135) uses the convexity of $\mathcal{L}(\boldsymbol{\beta}^{\text{lasso}})$. Denote event $\mathcal{E}_0$ as follows:

$$\mathcal{E}_0 = \left\{\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty < \frac{1}{2}\lambda\right\}. \tag{EC.137}$$

Under $\mathcal{E}_0$, (EC.136) can be further simplified into

$$
\begin{aligned}
-\frac{1}{2}\lambda\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1+\lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1 &\le \lambda\|\boldsymbol{\beta}^{\text{true}}\|_1\\
-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}\|_1 &\le \|\boldsymbol{\beta}^{\text{true}}\|_1\\
-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}\|_1 &\le \|\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1+\|\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1. \qquad \text{(EC.138)}
\end{aligned}
$$

As $\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}=\mathbf{0}$ by definition, we then have

$$
\begin{aligned}
&-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\mathbf{0}\|_1\le\|\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1+0\\
&-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1-\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}\|_1+\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1\le\|\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1+0\\
&\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1\le\frac{1}{2}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1+\|\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1-\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}\|_1\\
&\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}^c}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}^c}\|_1\le 3\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1 \qquad \text{(EC.139)}
\end{aligned}
$$

Then, by Lemma EC.1, we obtain

$$
\mathbb{P}\left((\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})^\top\nabla^2\mathcal{L}(\boldsymbol{\xi})(\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})\ge\frac{|\mathcal{A}|\kappa}{2ns}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1^2\right)\ge 1-\exp\left(-C_1|\mathcal{A}|\right). \qquad \text{(EC.140)}
$$

Now, we turn back to (EC.134) and use the Taylor expansion on $\mathcal{L}(\boldsymbol{\beta}^{\text{lasso}})$ at $\boldsymbol{\beta}^{\text{true}}$. Then, the following inequality holds for some $\boldsymbol{\xi}$ between $\boldsymbol{\beta}^{\text{true}}$ and $\boldsymbol{\beta}^{\text{lasso}}$:

$$
\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})+\frac{1}{2}(\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})^\top\nabla^2\mathcal{L}(\boldsymbol{\xi})(\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})+\lambda\|\boldsymbol{\beta}^{\text{lasso}}\|_1\le\lambda\|\boldsymbol{\beta}^{\text{true}}\|_1. \qquad \text{(EC.141)}
$$

Combining (EC.140) and (EC.141), we know that with probability at least $1-\exp(-C_1n)$, the following results hold:

$$
\begin{aligned}
\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1^2 &\le -\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})^\top(\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}})+\lambda\left(\|\boldsymbol{\beta}^{\text{true}}\|_1-\|\boldsymbol{\beta}^{\text{lasso}}\|_1\right)\\
\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\|_1^2 &\le \sum_{i\in\mathcal{S}\cup\mathcal{S}^c}\left[-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})(\beta_i^{\text{lasso}}-\beta_i^{\text{true}})+\lambda\left(|\beta_i^{\text{true}}|-|\beta_i^{\text{lasso}}|\right)\right]. \qquad \text{(EC.142)}
\end{aligned}
$$

We then separately consider $i\in\mathcal{S}$ and $i\in\mathcal{S}^c$ as follow:

$$
\begin{aligned}
&\sum_{i\in\mathcal{S}}\left[-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})(\beta_i^{\text{lasso}}-\beta_i^{\text{true}})-\lambda\left(|\beta_i^{\text{lasso}}|-|\beta_i^{\text{true}})|\right)\right]\\
&\le\left\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_\infty\left\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\right\|_1+\lambda\left\|\boldsymbol{\beta}^{\text{lasso}}_{\mathcal{S}}-\boldsymbol{\beta}^{\text{true}}_{\mathcal{S}}\right\|_1 \qquad \text{(EC.143)}
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{i\in\mathcal{S}^c}\left[-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})(\beta_i^{\text{lasso}}-\beta_i^{\text{true}})-\lambda\left(|\beta_i^{\text{lasso}}|-|\beta_i^{\text{true}})|\right)\right]\\
&\le\sum_{i\in\mathcal{S}^c}\left[-\nabla_i\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\beta_i^{\text{lasso}}-\lambda|\beta_i^{\text{lasso}}|\right]\\
&\le\sum_{i\in\mathcal{S}^c}\left(|\nabla_i\mathcal{L}(\beta^{\text{true}})|-\lambda\right)|\beta_i^{\text{lasso}}|\le 0, \qquad \text{(EC.144)}
\end{aligned}
$$

where the last inequality uses $\|\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\|_\infty\le\frac{1}{2}\lambda$ in $\mathcal{E}_0$.

Combining (EC.142), (EC.143) and (EC.144), we can show that

$$\begin{aligned}
&\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}}-\boldsymbol{\beta}_{S}^{\text{true}}\|_1^2 \leq \left\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_{\infty}\left\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}}-\boldsymbol{\beta}_{S}^{\text{true}}\right\|_1+\lambda\left\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}}-\boldsymbol{\beta}_{\mathcal{S}}^{\text{true}}\right\|_1\\
&\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}}-\boldsymbol{\beta}_{S}^{\text{true}}\|_1 \leq \left\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_{\infty}+\lambda\\
&\frac{|\mathcal{A}|\kappa}{4ns}\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1 \leq 4\left(\left\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_{\infty}+\lambda\right) && \text{(EC.145)}\\
&\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1 \leq \frac{16ns}{|\mathcal{A}|\kappa}\left(\left\|\nabla_{\mathcal{S}}\mathcal{L}(\boldsymbol{\beta}^{\text{true}})\right\|_{\infty}+\lambda\right), && \text{(EC.146)}
\end{aligned}$$

where (EC.145) uses $\|\boldsymbol{\beta}_{\mathcal{S}^c}^{\text{lasso}}-\boldsymbol{\beta}_{\mathcal{S}^c}^{\text{true}}\|_1 \leq 3\|\boldsymbol{\beta}_{\mathcal{S}}^{\text{lasso}}-\boldsymbol{\beta}_{\mathcal{S}}^{\text{true}}\|_1$ in (EC.139). Under event $\mathcal{E}_0$, (EC.146) can be further reduced to:

$$\|\boldsymbol{\beta}^{\text{lasso}}-\boldsymbol{\beta}^{\text{true}}\|_1 \leq \frac{24ns}{|\mathcal{A}|\kappa}\lambda. \tag{EC.147}$$

Now, we assess the probability of event $\mathcal{E}_0$. The i-th element of $\nabla\mathcal{L}(\boldsymbol{\beta}^{\text{true}})$ is $\frac{1}{n}\sum_{j=1}^{n}X_{j,i}f_y^{'}(R_j|\boldsymbol{X}_j^{\top}\boldsymbol{\beta}^{\text{true}})$. Under assumptions A.1 and A.4, $X_{j,i}f_y^{'}(R_j|\boldsymbol{X}_j^{\top}\boldsymbol{\beta}^{\text{true}})$ are $x_{\max}^2\sigma^2$-subguassian random variables for given sample $\boldsymbol{X}_j$. We can use the Hoeffding's inequality and union bound to build the following probability bound.

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{1}{n}\sum_{j=1}^{n}X_{j,i}f_y^{'}(R_j|\boldsymbol{X}_j^{\top}\boldsymbol{\beta}^{\text{true}})\right|\geq t\right)\leq 2\exp\left(-\frac{nt^2}{2\sigma^2x_{\max}^2}\right)\\
\Rightarrow&\mathbb{P}\left(\max_i\left|\frac{1}{n}\sum_{j=1}^{n}X_{j,i}f_y^{'}(R_j|\boldsymbol{X}_j^{\top}\boldsymbol{\beta}^{\text{true}})\right|\leq t\right)\geq 1-2d\exp\left(-\frac{nt^2}{2\sigma^2x_{\max}^2}\right), && \text{(EC.148)}
\end{aligned}$$

Setting $t=\frac{1}{2}\lambda$, we will have event $\mathcal{E}_0$ defined in (EC.137) holds with at least probability $1-2d\exp\left(-\frac{n\lambda^2}{8\sigma^2x_{\max}^2}\right)$. The desirable result directly follows by (EC.146), (EC.147), and (EC.148).

LEMMA EC.3. *Under assumptions A.1, A.3, A.4, and A.5, for any feasible $\boldsymbol{x}$, $\boldsymbol{\beta}_i$ and $i\in\mathcal{K}$, the following two statements hold.*

1. $|\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i^{\text{true}}]-\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i]|\leq R_{\max}e^{\sigma x_{\max}\|\boldsymbol{\beta}_i-\boldsymbol{\beta}_i^{\text{true}}\|_1}\sigma x_{\max}\|\boldsymbol{\beta}_i-\boldsymbol{\beta}_i^{\text{true}}\|_1$
2. *Moreover, if* $\|\boldsymbol{\beta}_i-\boldsymbol{\beta}_i^{\text{true}}\|_1\leq\min\left\{\frac{1}{\sigma x_{\max}},\frac{h}{4e\sigma R_{\max}x_{\max}}\right\}$, *then we have* $\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i]>\max_{j\neq i}\mathbb{E}_{\epsilon}[R_j|\boldsymbol{x}^{\top}\boldsymbol{\beta}_j]+\frac{h}{2}$ *for* $i\in\mathcal{K}_o$ *and* $\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i]<\max_{j\neq i}\mathbb{E}_{\epsilon}[R_j|\boldsymbol{x}^{\top}\boldsymbol{\beta}_j]-\frac{1}{2}h$ *for* $i\in\mathcal{K}_s$.

Proof of Lemma EC.3 ***To show the part 1.*** We first expand the left-hand-side as follows.

$$\begin{aligned}
&\left|\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i^{\text{true}}]-\mathbb{E}_{\epsilon}[R_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i]\right|\\
=&\left|\int_{-\infty}^{+\infty}r_ie^{-f(r_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i^{\text{true}})}dr_i-\int_{-\infty}^{+\infty}r_ie^{-f(r_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i)}dr_i\right| && \text{(EC.149)}\\
=&\left|\int_{-\infty}^{+\infty}r_i\left(e^{-f(r_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i^{\text{true}})}-e^{-f(r_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i)}\right)dr_i\right|.\\
=&\left|\int_{-\infty}^{+\infty}-r_i\left(e^{-f(r_i|\boldsymbol{x}^{\top}\boldsymbol{\beta}_i)}\right)'\Big|_{\boldsymbol{\beta}_i=\boldsymbol{\beta}_i^{\text{true}}+\boldsymbol{\delta}}\boldsymbol{x}^{\top}(\boldsymbol{\beta}_i-\boldsymbol{\beta}_i^{\text{true}})dr_i\right|, && \text{(EC.150)}
\end{aligned}$$

where (EC.149) uses f being the sample-wise negative log-likelihood loss function and $\boldsymbol{\delta}$ in (EC.150) is between $\mathbf{0}$ and $\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}$. We then pull $\boldsymbol{x}^T(\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}})$ out of the integral:

$$\begin{aligned}
&\left| \int_{-\infty}^{+\infty} -r_i \left(e^{-f(r_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i)} \right)' \Big|_{\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}} \boldsymbol{x}^\top (\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}) dr_i \right| \\
=& \left| \boldsymbol{x}^\top (\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}) \int_{-\infty}^{+\infty} -r_i \left(e^{-f(r_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i)} \right)' \Big|_{\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}} dr_i \right| \\
\leq& \left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}))} f_y'(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta})) dr_i \right| x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1. \qquad \text{(EC.151)}
\end{aligned}$$

As we assume $|f_y'(\cdot|\cdot)|$ is bounded by σ in assumption A.4, (EC.151) is upper bounded by

$$\begin{aligned}
&\left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}))} f_y'(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta})) dr_i \right| x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \\
\leq& \left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}))} dr_i \right| \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1. \qquad \text{(EC.152)}
\end{aligned}$$

We then expand term $f(r_i|\boldsymbol{x}^\top(\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}))$ in (EC.152), and there exists a $\boldsymbol{\xi}$ between $\boldsymbol{\beta}^{\text{true}}$ and $\boldsymbol{\beta}^{\text{true}} + \boldsymbol{\delta}$ such that

$$\begin{aligned}
&\left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top (\boldsymbol{\beta}_i^{\text{true}} + \boldsymbol{\delta}))} dr_i \right| \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \\
=& \left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}) - f_y'(r_i|\boldsymbol{x}^\top \boldsymbol{\xi}) \boldsymbol{x}^\top \boldsymbol{\delta}} dr_i \right| \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j^{\text{true}}\|_1 \\
\leq& \left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}) + |f_y'(r_i|\boldsymbol{x}^\top \boldsymbol{\xi})| \|\boldsymbol{x}\|_\infty \|\boldsymbol{\delta}\|_1} dr_i \right| \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \\
\leq& \left| \int_{-\infty}^{+\infty} r_i e^{-f(r_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}})} dr_i \right| e^{\sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1} \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \qquad \text{(EC.153)} \\
=& |\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}]| e^{\sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1} \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \qquad \text{(EC.154)}
\end{aligned}$$

where (EC.153) uses that $\boldsymbol{\delta}$ is between $\mathbf{0}$ and $\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}$, which implies $\|\boldsymbol{\delta}\|_1 \leq \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1$, and (EC.154) comes from the definition of $\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}]$. Combining $\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}] \in (0, R_{\max}]$ in assumption A.1 and (EC.154), we have:

$$\left| \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i] \right| \leq R_{\max} e^{\sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1} \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1. \qquad \text{(EC.155)}$$

To show the part 2. If $\|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \leq \frac{1}{\sigma x_{\max}}$, then we can show that

$$e^{\sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1} \leq e. \qquad \text{(EC.156)}$$

Combining (EC.156) and (EC.155), we obtain

$$\begin{aligned}
\left| \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i] \right| \leq& R_{\max} e^{\sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1} \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \\
\leq& R_{\max} e \sigma x_{\max} \|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \qquad \text{(EC.157)}
\end{aligned}$$

Let $j_1 = \arg\max_{j \neq i} \mathbb{E}_\epsilon[R_j|\boldsymbol{x}^\top \boldsymbol{\beta}_j^{\text{true}}]$. We first consider the case with $i \in \mathcal{K}_o$. Under assumption A.3, for any $x \in U_i$, $i \in \mathcal{K}_o$, the following inequalities hold:

$$\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top \boldsymbol{\beta}_i^{\text{true}}] > \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top \boldsymbol{\beta}_{j_1}^{\text{true}}] + h$$

$$
\begin{aligned}
&\Rightarrow \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] > \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}^{\text{true}}] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] \\
&\qquad + \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] + h \\
&\Rightarrow \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] > -\left|\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i^{\text{true}})]\right| \\
&\qquad - \left|\mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}^{\text{true}}] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}]\right| + h.
\end{aligned}
\tag{EC.158}
$$

If $\|\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}}\|_1 \le \frac{h}{4e\sigma R_{\max} x_{\max}}$, then we have

$$
\|R_{\max} e\sigma x_{\max}(\boldsymbol{\beta}_i - \boldsymbol{\beta}_i^{\text{true}})\|_1 \le \frac{h}{4}. \tag{EC.159}
$$

Combining (EC.157), (EC.158), and (EC.159), we will have

$$
\begin{aligned}
&\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] > -\frac{h}{4} - \frac{h}{4} + h \\
\Rightarrow &\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] > \max_{j\neq i} \mathbb{E}_\epsilon[R_j|\boldsymbol{x}^\top\boldsymbol{\beta}_j] + \frac{h}{2},
\end{aligned}
\tag{EC.160}
$$

where the last inequality uses $j_1 = \arg\max_{j\neq i} \mathbb{E}_\epsilon[R_j|\boldsymbol{x}^\top\boldsymbol{\beta}_j^{\text{true}}]$.

We then consider the case where $i \in \mathcal{K}_s$. Under assumption A.3, for any suboptimal arm $i \in \mathcal{K}_s$, we have

$$
\begin{aligned}
&\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i^{\text{true}}] + h < \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}^{\text{true}}] \\
\Rightarrow &\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] + h < \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}^{\text{true}}] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] \\
&\qquad + \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] \\
\Rightarrow &\mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] > -|\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i^{\text{true}}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i]| \\
&\qquad - |\mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}^{\text{true}}] - \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}]| + h \\
\Rightarrow &\mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] - \mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] > -\frac{1}{4}h - \frac{1}{4}h + h \\
\Rightarrow &\mathbb{E}_\epsilon[R_i|\boldsymbol{x}^\top\boldsymbol{\beta}_i] < \mathbb{E}_\epsilon[R_{j_1}|\boldsymbol{x}^\top\boldsymbol{\beta}_{j_1}] - \frac{1}{2}h = \max_{j\neq i} \mathbb{E}_\epsilon[R_j|\boldsymbol{x}^\top\boldsymbol{\beta}_j] - \frac{1}{2}h,
\end{aligned}
$$

where the second-to-last inequality uses (EC.157) and (EC.159).

Lemma EC.4. *Let α be a positive number. When $x > \max\{3\alpha^{-1}\log\alpha^{-1}, 0\}$, we have $\alpha x \ge \log x$.*

Proof of Lemma EC.4 Let $f(x) = \alpha x - \log x$. We first prove that for the case where $\alpha < e^{-1}$, $f(x)$ is non-negative for $x > \max\{3\alpha^{-1}\log\alpha^{-1}, 0\}$ via solving the following equation:

$$
\alpha x - \log x = 0 \Leftrightarrow \frac{\exp(\alpha x)}{x} = 1 \Leftrightarrow x\exp(-\alpha x) = 1 \Leftrightarrow -\alpha x \exp(-\alpha x) = -\alpha, \tag{EC.161}
$$

whose non-negative solution is $x = -\alpha^{-1}W_{-1}(-\alpha)$ where $W_{-1}(\cdot)$ is the Lambert W function. Combining this result with the monotonicity of $f(x)$, we have $f(x) \ge 0$ for all $x \ge -\alpha^{-1}W_{-1}(-\alpha)$. Next, we will show that $-\alpha^{-1}W_{-1}(-\alpha) \le \max\{3\alpha^{-1}\log\alpha^{-1}, 0\}$. By setting $u = -\log\alpha - 1$ in Theorem 1 of Chatzigeorgiou (2013), we have

$$
W_{-1}(-e^{-(-\log\alpha-1)-1}) \ge -1 - \sqrt{2(-\log\alpha - 1)} - (-\log\alpha - 1)
$$

$$\begin{aligned}
&\Rightarrow W_{-1}(-\alpha) \geq -\sqrt{2(-\log\alpha - 1)} + \log\alpha \\
&\Rightarrow W_{-1}(-\alpha) \geq -2\sqrt{-\log\alpha} + \log\alpha \\
&\Rightarrow -\alpha^{-1}W_{-1}(-\alpha) \leq \alpha^{-1}(2\sqrt{-\log\alpha} - \log\alpha) \\
&\Rightarrow -\alpha^{-1}W_{-1}(-\alpha) \leq -3\alpha^{-1}\log a = 3\alpha^{-1}\log\alpha^{-1} \leq \max\{3\alpha^{-1}\log\alpha^{-1}, 0\},
\end{aligned} \tag{EC.162}$$

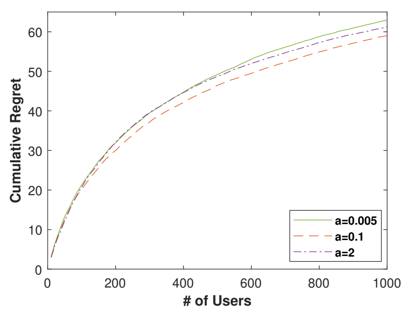
where the last inequality uses $-\log\alpha \geq \log e = 1$ when $\alpha \leq e^{-1}$. Therefore, we have $f(x) \geq 0$ for all $x \geq \max\{3\alpha^{-1}\log\alpha^{-1}, 0\}$.

Next, we consider the case where $\alpha \geq e^{-1}$. We can verify that $f(x)$ is convex with the minimum value $(1 + \log\alpha)$. If $\alpha \geq e^{-1}$, then $f(x)$ is non-negative, which implies that $\alpha x \geq \log x$ for all $x > 0$. Finally, the lemma follows directly by combining both the $\alpha < e^{-1}$ case and the $\alpha \geq e^{-1}$ case.

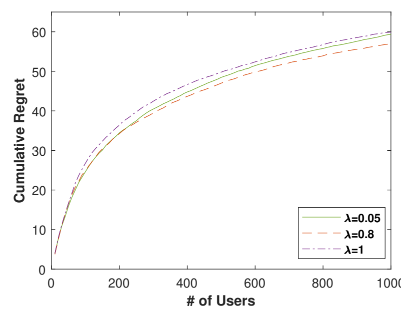
EC.3. Appendix: Sensitivity Analyses

In this section, we conduct sensitivity analyses for input parameters (i.e., a, $\lambda_{1,0}$, $\lambda_{2,0}$, h, t_0) in the G-MCP-Bandit algorithm and the upper-bound for significant covariates dimension (i.e., s). In particular, we will hold the baseline case, where $d=50$, $s=10$, $\lambda_1=\lambda_{2,0}=1$, $h=1$, $t_0=4$, $a=0.1$, and $k=2$, largely unchanged while varying one of $a\in\{0.005,0.1,2\}$, $\lambda=\lambda_1=\lambda_{2,0}\in\{0.05,0.8,1\}$, $h\in\{0.25,0.5,2\}$, $t_0\in\{0.3,0.5,1\}$, and $s\in\{3,4,5\}$. For each sensitivity analysis, we perform 100 trials and report the average cumulative regret for the G-MCP-Bandit algorithm up to 1000 users.

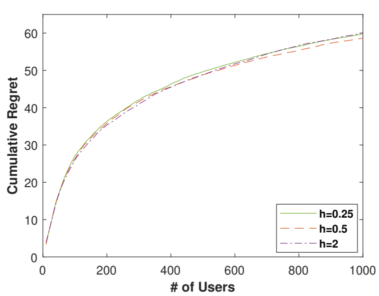
Figure EC.1 **Sensitivity analysis, where $d=50$, $s=10$, $\lambda_1=\lambda_{2,0}=1$, $h=1$, $t_0=4$, $a=0.1$, and $k=2$ as the baseline.**



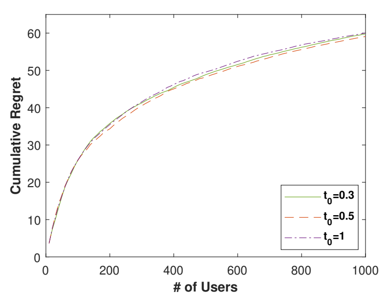
(a) Influence of a



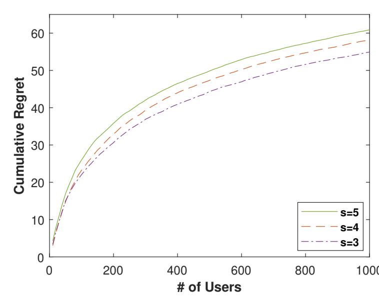
(b) Influence of $\lambda_1=\lambda_{2,0}=\lambda$



(c) Influence of h



(d) Influence of t_0



(e) Influence of s

We observe that the cumulative regret for the G-MCP-Bandit algorithm is robust with respect to the choices of its input parameters. Specifically, in Figure EC.1(a), (b), (c), and (d), the cumulative regret remains largely unchanged, when we vary the G-MCP-Bandit algorithm's input parameters (i.e., a, λ_1, $\lambda_{2,0}$, h, t_0). Furthermore, we find that the cumulative regret exhibits a non-monotonic pattern with respect to these input parameters changes and that the cumulative regret seems to be minimized for the median values of these parameters in EC.1(a), (b), (c), and (d). Hence, despite the mild improvement in the cumulative regret, actively tuning parameters may continue to be beneficial for decision-makers in practice.

At last, Figure EC.1(e) reports the influence of the upper bound for the significant covariates dimension (s) on the cumulative regret performance. In particular, we observe that the cumulative regret is monotonically increasing in s. Note that decreasing s suggests a higher sparsity level and a smaller number of significant covariates. Hence, as expected, with less non-zero parameters needed to be estimated, the G-MCP-Bandit algorithm will have better estimation accuracy, which in turn improves its regret performance.

EC.3.1. The Knowledge of the Sparsity Level s

To establish the G-MCP-Bandit algorithm's regret upper bound, some input parameters may need to be selected based on the sparsity level s. For example, the parameter a is chosen to ensure the condition $a > \frac{144s}{\kappa}$ holds in Proposition 1 and Theorem 1. Such a selection condition is standard in the high-dimensional statistics literature (e.g., Corollary 4 and Corollary 6 of Fan et al. 2014b and Lemma 5.3 of Wang et al. 2014) and the high-dimensional bandit literature (e.g., Proposition 1 of Bastani and Bayati 2020). In practice, however, decision-makers may not know the sparsity level s, especially without sufficient data at the beginning. Therefore, in this subsection, we will investigate the question of if decision-makers don't know the sparsity level s, then how the suboptimal parameter selection will influence G-MCP-Bandit's regret performance.

In particular, we use $\hat{s}$ to represent decision-makers' guess or estimation of the true sparsity level s. Without knowing the true s value, decision-makers will tune the G-MCP-Bandit algorithm's parameters by using their estimated sparsity level $\hat{s}$. In Figure EC.2, we report five linear two-armed bandit experiments[5], in which the covariate dimension $d = 100$ and the true sparsity level $s \in \{5, 20, 30, 40, 50\}$. For each experiment, we perform 30 trials and report the average cumulative regret of five G-MCP-Bandit algorithms[6] that are tuned by using/assuming $\hat{s} = 5$, 20, 30, 40, and 50, respectively[7]. Therefore, in each experiment, only one G-MCP-Bandit algorithm's parameters are tuned by the true sparsity level s, and the other remaining four G-MCP-Bandit algorithms used the suboptimal parameters tuned by the wrong estimated sparsity level $\hat{s}$.

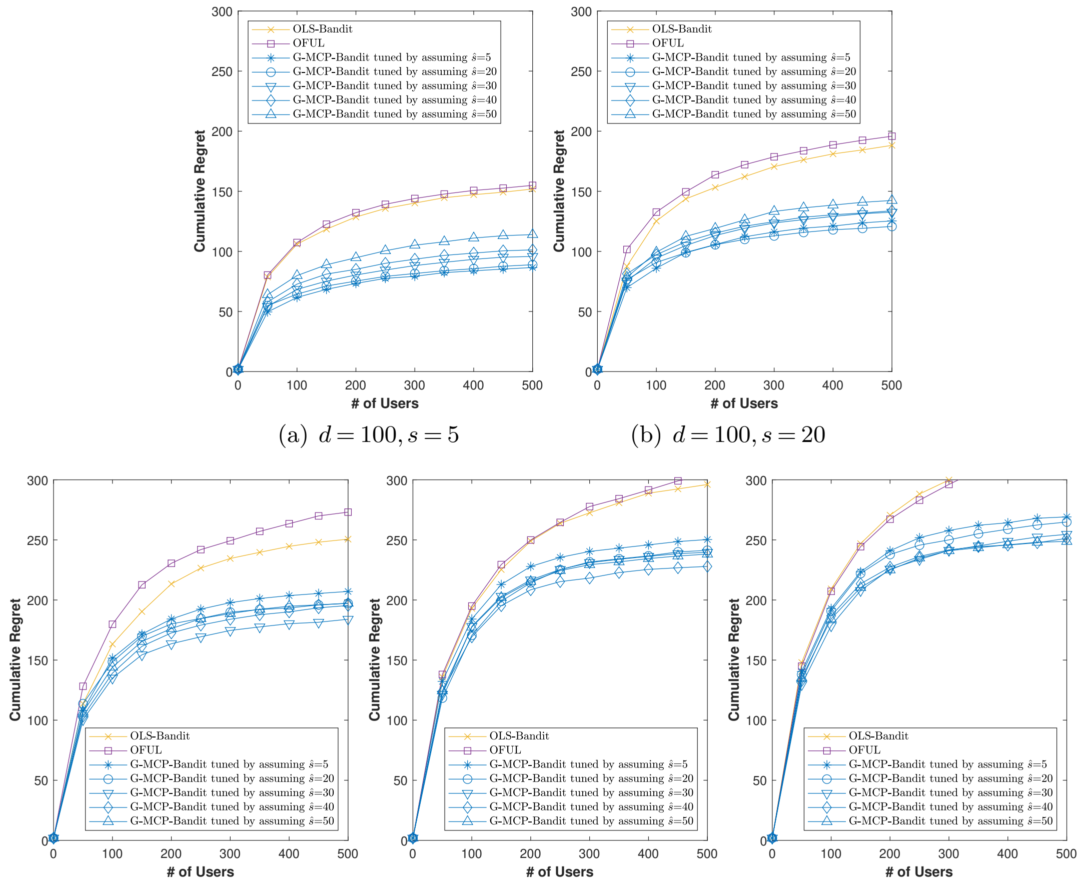
We observe that the G-MCP-Bandit algorithm's cumulative regret will be minimized when it is tuned by using the correct sparsity value (i.e., if $\hat{s} = s$). For example, in Figure EC.2(a), where the true sparsity level $s = 5$, the G-MCP-Bandit algorithm will have the lowest cumulative regret, if it was tuned by using $\hat{s} = 5$ (see the blue line with asterisk marks). In addition, we find that the larger the distance between $\hat{s}$ and s, the worse the G-MCP-Bandit algorithm will perform. And, the regret differences among the G-MCP-Bandit algorithms tuned by different $\hat{s}$ values tend to be narrowed as the true sparsity level s increases. The regret differences among G-MCP-Bandit algorithms tuned by different $\hat{s}$ values highlight the importance of accurately estimating s. Therefore, to improve the G-MCP-Bandit algorithm's performance, decision-makers should, whenever possible, (1) use their earlier experience or data previously obtained from similar scenarios to improve the accuracy of estimating the sparsity level and (2) dynamically adjust the tuning parameters for the G-MCP-Bandit algorithm, when more data become available to support a better estimation of the s value.

[5] Other experiments exhibit similar pattern and therefore are omitted.

[6] The Lasso-Bandit algorithm also requires the knowledge of the sparsity level to tune parameters, and therefore its cumulative regret performance depends on the estimated sparsity level $\hat{s}$. In our experiments, we observe that the G-MCP-Bandit algorithm continues to outperform the Lasso-Bandit algorithm, if both were tuned under the same $\hat{s}$. In addition, as the impact of the estimated sparsity level $\hat{s}$ on the Lasso-Bandit algorithm is nearly identical to that on the G-MCP-Bandit algorithm, we omit the Lasso-Bandit algorithm for better clarity in the figure.

[7] We also tried $\hat{s}$ value that is higher than 50, but the cumulative regret performance for cases with $\hat{s} > 50$ is very close to the $\hat{s} = 50$ case and therefore will be omitted to avoid duplication.

Figure EC.2 The influence of unknown s, where parameters for the G-MCP-Bandit algorithm are optimally tuned by assuming decision-makers' estimated sparsity levels to be $\hat{s} = 5, 20, 30, 40,$ and 50.



(a) $d = 100, s = 5$ (b) $d = 100, s = 20$

(c) $d = 100, s = 30$ (d) $d = 100, s = 40$ (e) $d = 100, s = 50$

Remark EC.1. To theoretically remove the dependence of parameter a on the sparsity level s, we will need to revise the existing assumption for a stronger version or introduce new assumptions. The dependence of a on s comes from the proof of Proposition 3, in which we want to ensure the penalty weight w_i for $i \in \mathcal{S}^c$ to be positive. According to (EC.16), we must set $a = \mathcal{O}(s)$. One way to break such a dependence is to further separate $\mathcal{S}^c$ into two subsets $\mathcal{S}_3$ and $\mathcal{S}_4$. We then count the index in $\mathcal{S}^c$ with small enough penalty weight in $\mathcal{S}_3$ and $\mathcal{S}_4 = \mathcal{S}^c / \mathcal{S}_3$. As the element in $\mathcal{S}_3$ indicates that the magnitude of the Lasso estimator is large while the ground truth is 0, it will happen with low probability. Thus, the cardinality of $\mathcal{S}_3$ will not be large. In fact, we can prove that $|\mathcal{S}_3| \le s$ under some additional mild conditions. Then, by setting $\hat{\mathcal{S}} = \mathcal{S} \cup \mathcal{S}^3$ and $\hat{\mathcal{S}}^c = \mathcal{S}^4$, for all i in $\hat{\mathcal{S}}^c$, the penalty weights will be large enough. Therefore, if we can further introduce a stronger restricted eigenvalue condition so that $\frac{\kappa}{s}\|\boldsymbol{u}_{\hat{\mathcal{S}}}\|_1^2 \le \boldsymbol{u}^\top \mathbb{E}[\nabla^2 \mathcal{L}(\boldsymbol{\xi})]\boldsymbol{u}$ holds for the index set $\hat{\mathcal{S}}$ with $|\hat{\mathcal{S}}| \le 2s$, the proof of Proposition 3 will be able to be established without a dependence on s.

The knowledge of the s value in λ_1 and C_0 can be resolved by replacing s with $\hat{s}\sqrt{\log t}$, where $\hat{s}$ is a guess or estimation on s. In our setting, s is defined as an upper bound for the cardinality of the significant index sets for all arms, so our analysis works for the setting with an over-estimation on s. If we set $s = \hat{s}\sqrt{\log t}$

in the algorithm, for a large enough t, we will enter the over-estimating scenario and be able to recover the desired statistical properties of our algorithm, even when the initial parameter $\hat{s}$ is incorrectly specified to be small. However, the regret during the initial time periods may suffer as a result. We exclude the proof for brevity.

We can also remove the dependence on s by introducing additional assumptions on covariates diversity/balance from the nearly exploration-free bandit literature (e.g., Assumption 3 in Bastani et al. 2021 and Assumption 6 in Oh et al. 2021). With these assumptions, we can directly ensure that Proposition 6 holds even without enough random samples or with a wrong construction for the optimal decision set Π_t in the G-MCP-Bandit algorithm. The intuition is that by introducing these assumptions, we can ensure that all dimensions are explored with a nearly equal chance so that the MCP estimator under 2sWL will have a high probability to reach $\mathcal{S}_1 = \mathcal{S}$. Therefore, even if we use wrong C_0 and λ_1 from under-estimating s, the desirable regret bounds will continue to hold.

EC.4. Appendix: Additional Experiments on Tencent dataset

In this section, we extend the Tencent search advertising experiments in §6.2 by considering the impacts of a large number of ads and the robustness of the G-MCP-bandit algorithm under the model misspecification, where the underlying reward function is not within the family of GLMs.

EC.4.1. The impact of a large number of ads

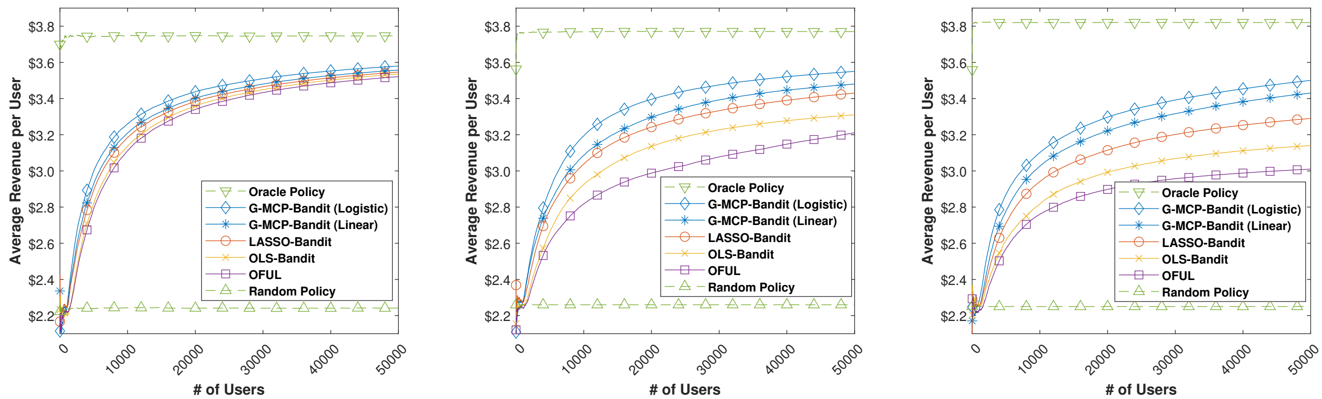
In this subsection, we expand the Tencent search advertising experiment to understand the impacts of a large number of ads. To be able to accurately estimate the true parameter vectors for the oracle policy, it is necessary to include ads with large session entries in all experiments. Hence, we first rank all ads that have CTR higher than 1% by their frequencies and then pick the top 10, 100, and 1000 ads for three experiments. The ads with the lowest frequency in these three experiments (i.e., $K = 10$, $K = 100$, and $K = 1000$) have 188997, 28954, and 2235 session entries, respectively, to provide estimations for parameter vectors under the oracle policy with reasonable accuracy. The reward for each clicked ad is initialized at the beginning of each experiment and randomly assigned to be \$1, \$5, or \$10 with equal probability.

First, as expected, we observe that the computational time increases in the number of ads and the number of users. In particular, with the Intel Xeon Platinum 8163 CPU (2.50GHz, 7 cores), the average computational time (in seconds) for the G-MCP-Bandit algorithm to complete $20,000$ users is 203 for $K = 10$, 226 for $K = 100$, and 349 for $K = 1000$. When the number of users is increased to $40,000$, the average computational time will increase to 583, 661, and 1375 seconds, respectively. Similar to §6.2, we benchmark the G-MCP-Bandit algorithm to OFUL, OLS-Bandit, Lasso-Bandit, the random policy, and the oracle policy. For each experiment, we perform 10 trials for each algorithm and report the average revenue with up to $50,000$ users.

Similar to the three-ad experiment in §6.2, we observe that the G-MCP-Bandit algorithm outperforms other algorithms in terms of the average revenue performance; see Figure EC.3. When the number of ads is comparatively small (e.g., $K = 10$), it does not need many users for all algorithms to identify the significant covariates and/or to estimate parameter vectors to eventually select the optimal ads for incoming users. Hence, the revenue improvement of the G-MCP-Bandit algorithm over other algorithms is most significant when the number of users is not too large. For example, when $T < 10000$, the revenue improvement of the G-MCP-Bandit algorithm over Lasso-Bandit is around $3\% - 4\%$. As the number of users increases, all algorithms eventually learn to accurately estimate parameter vectors to identify the revenue-maximizing ads. Therefore, the average revenue performance of all algorithms begins to converge.

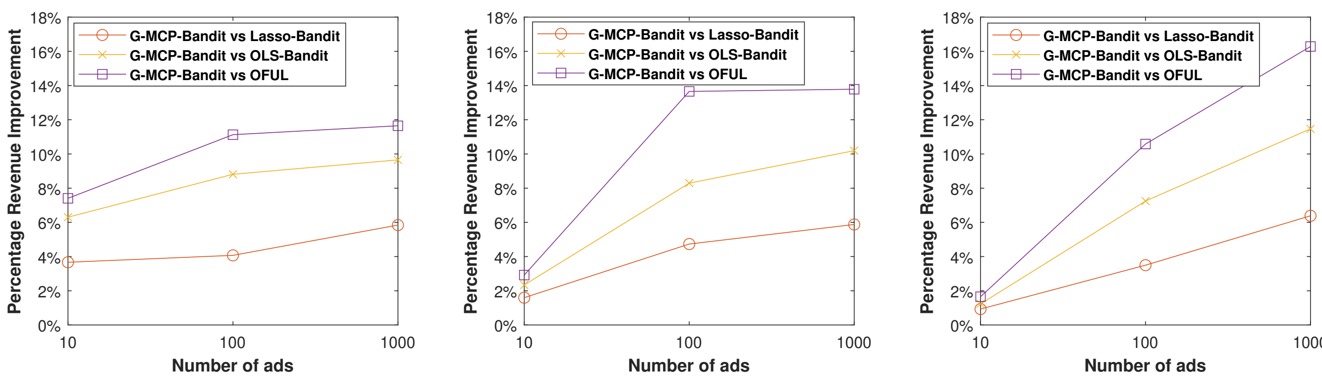
As the number of ads increases (e.g., $K = 100$ and $K = 1000$), accurately learning the parameter vectors and identifying the optimal ad require more users, which is where the G-MCP-Bandit algorithm shines the most. In particular, we observe that the revenue improvement of the G-MCP-Bandit algorithm over other algorithms tends to grow with the number of ads. Figure EC.4 plots the percentage revenue improvement of the G-MCP-Bandit algorithm over other benchmarking algorithms by fixing $T = 5000$, $T = 20000$, and $T = 50000$. In all three scenarios, the benefits of the G-MCP-Bandit algorithm increase with the number of ads. This observation suggests that the G-MCP-Bandit algorithm becomes more favorable in practice, especially when there are large pools of available ads for decision-makers to choose from.

Figure EC.3 The impact of the number of ads K on average revenue.



(a) K=10 (b) K=100 (c) K=1000

Figure EC.4 The percentage revenue improvement of G-MCP-Bandit (Logistic) over other algorithms.



(a) T=5000 (b) T=20,000 (c) T=50,000

EC.4.2. Robustness under model misspecification

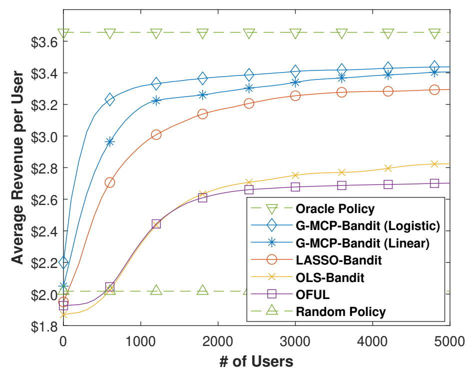
In §6.2 and Appendix EC.4.1, we have examined the robustness of the G-MCP-Bandit algorithm under the model misspecification, where all algorithms assume the linear model, but the true underlying reward function actually follows the logistic model. Under such a misspecified setting, the G-MCP-Bandit algorithm under the linear model outperforms all other algorithms in terms of average revenue performance.

In this subsection, we conduct another experiment to further test the robustness of the G-MCP-Bandit algorithm under the model misspecification, where the true underlying model does not belong to the GLMs family. In particular, we consider the scenario where the true underlying model follows the form of a two-component Gaussian Mixture Model (GMM), which does not belong to the GLMs family. Theoretically, GMM has better representation power than GLMs, and for the Tencent dataset, it actually fits the Tencent data better[8] than both the linear model and the logistic model. Analogous to Figure 3 in the main paper,

[8] We train the GMM model with around three hundred covariates with the highest frequency.

we consider the same three-ad experiment. For each algorithm, we perform 10 trials and report the average revenue with up to 5000 users in Figure EC.5.

Figure EC.5 The robustness of the G-MCP-Bandit algorithm under the model misspecification, where the true underlying model follows a two-component Gaussian Mixture Model.



Consistent with all previous experiments, Figure EC.5 shows that the G-MCP-Bandit algorithm, under both the linear model and the logistic model, continues to outperform other algorithms in terms of average revenue performance. In addition, we observe that all algorithms in Figure EC.5 seem to generate less average revenue that what is shown in Figure 3. This observation may be due to the fact that it is much more difficult to use a GLM model (e.g., a linear or logistic model) to approximate the GMM model than to use the linear model to approximate the logistic model, so the impacts of the model misspecification on the average revenue performance are much more severe in Figure EC.5 than in Figure 3. Despite the negative impacts of the model misspecification, the G-MCP-Bandit algorithm continues to outperform Lasso-Bandit by 7.30% (under the logistic model) and 4.21% (under the linear model), and such an improvement is even larger when compared to OLS-Bandit and OFUL-Bandit.