Supplement to “Online Batch Decision-Making with High-Dimensional Covariates”

A Proof of Theorem 1

Proof of Theorem 1. Based on standard arguments in high dimensional statistics, the Template LASSO on $\mathcal{A}_{w_k}$, when choosing $\lambda \geq 2\lambda_0(\gamma)$, satisfies

$$P(\|\widehat{\beta}(\mathcal{A}_{w_k}, \lambda) - \beta_{w_k}\|_1 \leq 4\lambda \frac{|S_{w_k}|}{\phi^2}) \geq P[\mathcal{F}(\lambda_0(\gamma))] - P(\widehat{\Sigma}(\mathcal{A}_{w_k}) \notin C(S_{w_k}, \phi))) \tag{20}$$

where $\mathcal{F}(\lambda_0(\gamma)) \equiv \left\{\max_{1\leq j\leq p} \frac{2}{T} \left|\eta^\top X^{(j)}\right| \leq \lambda_0(\gamma)\right\}$, which is a high-probability event by carefully choosing the threshold $\lambda_0(\gamma)$ stated in the following lemma:

Lemma 2. *Given a sample set $\mathcal{A}_{w_k}$ and choose $\lambda_0(\gamma) \equiv 2\sigma x_{\max}\sqrt{(\frac{\gamma^2+2\log p)}{|\mathcal{A}_{w_k}|}}$, then*

$$P(\mathcal{F}(\lambda_0(\gamma))) \geq 1 - 2\exp[-\frac{\gamma^2}{2}] \tag{21}$$

Proof. See Section B.1. ◻

Besides, the sample covariance matrix of the template sample set $\mathcal{A}_{w_k}$ satisfies the compatibility condition with high probability, as stated in the following lemma:

Lemma 3. *Given a sample set $\mathcal{A}_{w_k}$ that satisfies rate r optimal allocation condition. Then*

$$P(\widehat{\Sigma}(\mathcal{A}_{w_k}) \notin \mathcal{C}(\beta_{S_{w_k}}, \frac{\phi}{\sqrt{2}}\sqrt{\frac{|\mathcal{A}^\sharp_{w_k}|}{|\mathcal{A}_{w_k}|}})) \leq \exp(-C_2(\phi_1)^2|\mathcal{A}^\sharp_{w_k}|). \tag{22}$$

Proof. See Section B.2. ◻

Now, lemma 2 and lemma 3 together turn the equation (20) into

$$P(\|\widehat{\beta}(\mathcal{A}_{w_k}, \lambda) - \beta_{w_k}\|_1 \leq 4\lambda \frac{|S_{w_k}|}{\phi^2}) \geq 1 - 2\exp[-\frac{\gamma^2}{2}] - \exp(-C_2(\phi_1)^2|\mathcal{A}^\sharp_{w_k}|).$$

Then Theorem 1 follows by solving γ from the condition $\lambda \geq 2\lambda_0(\gamma)$.

B Proof of key Lemmas

B.1 Proof of Lemma 2

Lemma 2 The event

$$\mathcal{F}(\lambda_0(\gamma)) = \{\max_{j\in[p]} \frac{2}{T}|\eta^\top X^{(j)}| \leq \lambda_0(\gamma)\}. \tag{23}$$

holds with probability at least $1 - \exp(-\frac{\gamma^2}{2})$ by choosing

$$\lambda_0(\gamma) = 2x_{\max}\sigma\sqrt{\frac{\gamma^2 + 2\log d}{\sum_{s=1}^T n_{s,w_k}}}. \tag{24}$$

Proof. Let n_{s,w_k} denote the number of users allocated to treatment w_k at the sth decision epoch. The sample collected at epoch s is denoted by $\mathcal{A}_{s,k} = \{((X_{(i,s)}, Y_{(i,s)}) : i \in [n_{t,w_k}], s \in [t])\}$. Recall $X^{(j)}$ is the jth column of covariate matrix X and the good event

$$\mathcal{F}(\lambda_0(\gamma)) = \{\max_{j\in[p]} \frac{2}{T}|\eta^\top X^{(j)}| \leq \lambda_0(\gamma)\}. \tag{25}$$

With the help of union bound, we have

$$P\left(\mathcal{J}\left(\lambda_0(\gamma)\right)\right) \geq 1-\sum_{j=1}^{p} P\left(\left|\eta^\top X^{(j)}\right| > \frac{T}{2}\lambda_0(\gamma)\right).$$

The quantity $\eta^\top X^{(j)}$ actually has sub-Gaussian tail. To see this, define the filtration

$$F_t \equiv \{(Y_{(i,j)}, X_{(i,j)})\}_{i\in[t-1], j\in[n_{(t-1),w_k}]} \quad \text{and} \quad \mathcal{G}_t = \mathcal{F}_t \cup \{X_{(s,i)}\}_{i\in[n_{t,\omega_k}]} \tag{26}$$

From the tower property of expectation, independence between $\{\eta_{(s,i)}\}_{i\in[n_{s,w_k}]}$ given $\mathcal{G}_s$, sub-Gaussian assumption on $\eta_{(s,i)}$, bounded assumption on covariates $\left\|X_{(i,t)}\right\|_\infty \leq x_{max}$,we have

$$\begin{aligned}
& E[\exp(u\sum_{s=1}^{n_{s,\omega_k}} \eta_{(s,i)}X_{(s,i),j})|\mathcal{F}_s] & (27)\\
= \;& E[E[\exp(u\sum_{s=1}^{n_{s,\omega_k}} \eta_{(s,i)}X_{(s,i),j})|\mathcal{G}_s]|\mathcal{F}_s] & (28)\\
= \;& E[\prod_{s=1}^{n_{s,\omega_k}} E[\exp(u\eta_{(s,i)}X_{(s,i),j})|\mathcal{G}_s]|\mathcal{F}_s] & (29)\\
\leq \;& E[\prod_{s=1}^{n_{s,\omega_k}} \exp(u^2\frac{(\sigma X_{(s,i),j})^2}{2})|\mathcal{F}_s] \leq \prod_{s=1}^{n_{s,\omega_k}} \exp(u^2\frac{(\sigma x_{\max})^2}{2}) & (30)\\
= \;& \exp(u^2\frac{(\sqrt{n_{s,\omega_k}}\sigma x_{\max})^2}{2}). & (31)
\end{aligned}$$

The above result gives us a bound on the moment generating function of $\eta^\top X^{(j)}$ that

$$\begin{aligned}
E[\exp(u(\eta^\top X^{(j)}))] &= E[\exp(u\sum_{t=1}^{T}\sum_{i=1}^{n_{t,\omega_k}} \eta_{(s,i)}X_{(s,i),j})] & (32)\\
&\leq \exp(u^2\frac{(\sqrt{n_{T,\omega_k}}x_{\max}\sigma)^2}{2})E[\exp(u\sum_{t=1}^{T-1}\sum_{i=1}^{n_{t,\omega_k}} \eta_{(s,i)}X_{(s,i),j})] & (33)\\
&\leq \cdots \leq \exp(u^2\frac{(\sqrt{\sum_{s=1}^{T} n_{s,\omega_k}}x_{\max}\sigma)^2}{2}). & (34)
\end{aligned}$$

We find $\eta^\top X^{(j)}$ is $(\sqrt{\sum_{s=1}^{T} n_{s,\omega_k}}x_{\max}\sigma)^2$-sub-Gaussian. The tail probability bound of sub-Gaussian distribution gives

$$P\left(\left|\eta^\top X^{(j)}\right| > t\right) \leq 2\exp\left(-\frac{t^2}{2\sum_{s=1}^{T} n_{s,w_k}x_{\max}^2\sigma^2}\right).$$

Now, to reformat this into a desired tail probability form, we note

$$\begin{aligned}
1-2\exp(-\frac{\gamma^2}{2}) &= 1-\sum_{j=1}^{d} P(|\eta^\top X^{(j)}| > \frac{\sum_{s=1}^{T} n_{s,w_k}}{2}\lambda_0(\gamma)) & (35)\\
&\geq 1-2\exp(-\frac{\sum_{s=1}^{T} n_{s,w_k}\lambda_0^2(\gamma)}{8x_{\max}^2\sigma^2}+\log d). & (36)
\end{aligned}$$

The above suggests us to choose

$$\lambda_0(\gamma) = 2x_{\max}\sigma\sqrt{\frac{\gamma^2+2\log d}{\sum_{s=1}^{T} n_{s,w_k}}}.$$

□

B.2 Proof of Lemma 3

Lemma 3 Given a sample set $\mathcal{A}_{w_k}$ satisfying template condition with rate r. Then

$$P(\widehat{\Sigma}(\mathcal{A}_{w_k}) \notin \mathcal{C}(\beta_{S_{w_k}}, \frac{\phi}{\sqrt{2}}\sqrt{\frac{|\mathcal{A}^{\sharp}_{w_k}|}{|\mathcal{A}_{w_k}|}})) \leq \exp(-C_2(\phi_1)^2|\mathcal{A}^{\sharp}_{w_k}|). \tag{37}$$

Proof. From our population assumption, the population covariance matrix Σ_{w_k} satisfies the compatability condition $\Sigma_{w_k} \in \mathcal{C}(\beta_{S_{w_k}}, \phi)$. By carefully controlling $|\mathcal{A}_{w_k}|$ and $|\mathcal{A}^{\sharp}_{w_k}|/|\mathcal{A}_{w_k}|$, one could first show $\|\Sigma_{w_k} - \widehat{\Sigma}(\mathcal{A}^{\sharp}_{w_k})\|_\infty \leq \frac{\phi^2}{32|S_{w_k}|}$, which implies $\widehat{\Sigma}(\mathcal{A}^{\sharp}_{w_k}) \in \mathcal{C}(\beta_{S_{w_k}}, \frac{\phi}{\sqrt{2}})$ with high probability (by using Corollary 6.8 in page 152 of Bühlmann and Van De Geer (2011)). Next, by estimating an upper bound of the quadratic form induced by the covariance matrix of sample set $\mathcal{A}_{w_k}$, we can show, with high probability, that

$$\widehat{\Sigma}(\mathcal{A}_{w_k}) \in \mathcal{C}(\beta_{S_{w_k}}, \frac{\phi}{\sqrt{2}}\sqrt{\frac{|\mathcal{A}^{\sharp}_{w_k}|}{|\mathcal{A}_{w_k}|}}).$$

□

C Theory of LASSO

C.1 Basic Inequality

Lemma 4. *(Basic Ineuality from Optimality Condition) In LASSO,*

$$\frac{1}{n}\|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda\|\hat{\beta}\|_1 \leq \frac{2}{n}\epsilon^\top X(\hat{\beta} - \beta^0) + \lambda\|\beta^0\|_1. \tag{38}$$

Proof. To perform optimality analysis, we play with true beta β^0 and empirical minimizer $\hat{\beta}$(Short hand of $\widehat{\beta}_{w_k}(\mathcal{A}_k, \lambda)$). From the argument min, we start with

$$\frac{1}{n}\|Y - X\hat{\beta}\|_2^2 + \lambda\|\hat{\beta}\|_1 \leq \frac{1}{n}\|Y - X\beta^0\|_2^2 + \lambda\|\beta^0\|_1 \tag{39}$$

Direct calculation gives us

$$\begin{aligned}
& \frac{1}{n}\|Y - X\hat{\beta}\|_2^2 - \frac{1}{n}\|Y - X\beta^0\|_2^2 && (40)\\
= \;& \frac{1}{n}[Y^\top Y - 2Y^\top X\hat{\beta} + \hat{\beta}^\top X^\top X\hat{\beta}] - \frac{1}{n}[Y^\top Y - 2Y^\top X\beta^0 + (\beta^0)^\top X^\top X\beta^0] && (41)\\
= \;& \frac{1}{n}[2Y^\top X(\beta^0 - \hat{\beta}) + \hat{\beta}^\top X^\top X\hat{\beta} - (\beta^0)^\top X^\top X\beta^0] && (42)\\
= \;& \frac{1}{n}[2(X\beta^0 + \epsilon)^\top X(\beta^0 - \hat{\beta}) + \hat{\beta}^\top X^\top X\hat{\beta} - (\beta^0)^\top X^\top X\beta^0] && (43)\\
= \;& \frac{1}{n}[2(\beta^0)^\top X^\top X(\beta^0 - \hat{\beta}) + 2\epsilon^\top X(\beta^0 - \hat{\beta}) + \hat{\beta}^\top X^\top X\hat{\beta} - (\beta^0)^\top X^\top X\beta^0] && (44)\\
= \;& \frac{1}{n}[2\epsilon^\top X(\beta^0 - \hat{\beta}) + (\beta^0)^\top X^\top X(\beta^0) - 2(\beta^0)^\top X^\top X\hat{\beta} + \hat{\beta}^\top X^\top X\hat{\beta}] && (45)\\
= \;& \frac{1}{n}[2\epsilon^\top X(\beta^0 - \hat{\beta}) + (\beta^0 - \hat{\beta})X^\top X(\beta^0 - \hat{\beta})] && (46)\\
= \;& \frac{2}{n}\epsilon^\top X(\beta^0 - \hat{\beta}) + \frac{1}{n}\|X(\hat{\beta} - \beta^0)\|_2^2 && (47)
\end{aligned}$$

□

Lemma 5. *(Basic Inequality on Good Event) On good event $\mathcal{F}$ and with $\lambda \geq 2\lambda_0$, the basic inequality can be further reduced to*

$$\frac{2}{n}\|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda\|\hat{\beta}_{S_0^c}\|_1 \leq 3\lambda\|\hat{\beta}_{S_0} - \beta^0_{S_0}\|_1 \tag{48}$$

Proof. Recall the basic inequality

$$\frac{1}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+\lambda\|\hat{\beta}\|_1 \leq \frac{2}{n}\epsilon^\top X(\hat{\beta}-\beta^0)+\lambda\|\beta^0\|_1$$

Multiply it by 2 to get

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda\|\hat{\beta}\|_1 \leq 2\cdot\frac{2}{n}\epsilon^\top X(\hat{\beta}-\beta^0)+2\lambda\|\beta^0\|_1 \tag{49}$$

Plug in the upper bound to get

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda\|\hat{\beta}\|_1 \leq 2\cdot(\max_{j\in[p]}\frac{2}{n}|\epsilon^\top X^{(j)}|)\|\hat{\beta}-\beta^0\|_1+2\lambda\|\beta^0\|_1 \tag{50}$$

Then on good event $\mathcal{F}$, it becomes

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda\|\hat{\beta}\|_1 \leq 2\cdot\lambda_0\|\hat{\beta}-\beta^0\|_1+2\lambda\|\beta^0\|_1 \tag{51}$$

Apply $\lambda \geq 2\lambda_0$ to get

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda\|\hat{\beta}\|_1 \leq \lambda\|\hat{\beta}-\beta^0\|_1+2\lambda\|\beta^0\|_1 \tag{52}$$

To further reduce equation (52), we play with sparsity component. Let S_0 denote the sparsity location of truth β^0.

One the RHS, since $\beta^0_{S_0^c}=0$, we have an identity

$$\|\hat{\beta}-\beta^0\|_1=\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1+\|\hat{\beta}_{S_0^c}\|_1 \tag{53}$$

On the LHS, we have identity

$$\|\beta^0\|_1=\|\beta^0_{S_0}\|_1+\|\beta^0_{S_c}\|_1=\|\beta^0_{S_0}\|_1 \tag{54}$$

On the other hand, the empirical minimizer $\hat{\beta}$ only has identity

$$\|\hat{\beta}\|_1=\|\hat{\beta}_{S_0}\|_1+\|\hat{\beta}_{S_c}\|_1 \tag{55}$$

To link $\hat{\beta}_{S_0}$ with $\beta^0_{S_0}$ in L_1 norm, the inverse triangle inequality gives

$$\|\hat{\beta}_{S_0}\|_1 \geq \|\beta^0_{S_0}\|_1-\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1 \tag{56}$$

Then we have an inequality

$$\|\hat{\beta}\|_1 \geq \|\beta^0_{S_0}\|_1-\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1+\|\hat{\beta}_{S_0^c}\|_1 \tag{57}$$

Combine these two observations into equation (52), we have inequality

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda(\|\beta^0_{S_0}\|_1-\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1+\|\hat{\beta}_{S_0^c}\|_1) \leq \lambda(\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1+\|\hat{\beta}_{S_0^c}\|_1)+2\lambda\|\beta^0_{S_0}\|_1 \tag{58}$$

Reorganize them into the inequality

$$\frac{2}{n}\|X(\hat{\beta}-\beta^0)\|_2^2+2\lambda\|\hat{\beta}_{S_0^c}\|_1 \leq 3\lambda\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1+\lambda\|\hat{\beta}_{S_0^c}\|_1 \tag{59}$$

□

Lemma 6. *(Compatibility passes L_1 norm to square root of L_2 norm)*

On good event $\mathcal{F}$, $\lambda \geq 2\lambda_0$, and compatability condition associate with gram matrix $\hat{\Sigma}$ holds,

$$\|\hat{\beta}_{S_0}-\beta^0_{S_0}\|_1 \leq \frac{\sqrt{s_0}}{\phi_0}\frac{1}{\sqrt{n}}\|X(\hat{\beta}-\beta^0)\|_2 \tag{60}$$

Proof. To further reduce the basic inequality on good event (48), we impose condition on sparsity component S_0. In lemma 5, two quantities we play with are $\|\hat{\beta}_{S_0^c}\|_1$ and $\|\hat{\beta}_{S_0} - \beta^0_{S_0}\|_1$.

An implication of equation (48) is, on good event $\mathcal{F}$, it is true that

$$\|\hat{\beta}_{S_0^c} - \beta^0_{S_0^c}\|_1 = \|\hat{\beta}_{S_0^c}\|_1 \leq 3\|\hat{\beta}_{S_0} - \beta^0_{S_0}\|_1, \tag{61}$$

that is, on good event $\mathcal{F}$, the discrepancy between empirical minimizor and truth $\hat{\beta} - \beta^0$ always belongs to the class

$$\{\beta | \|\beta_{S_0^c}\|_1 \leq 3\|\beta_{S_0 \|_1}\}. \tag{62}$$

On such class, the **compatability condition** with Gram matrix $\hat{\Sigma} \equiv \frac{1}{n} X^\top X$ is

$$\|\beta_{S_0}\|_1 \leq \frac{\sqrt{s_0}}{\phi_0} \sqrt{\beta^\top \hat{\Sigma} \beta} \tag{63}$$

Thus, we have

$$\|\hat{\beta}_{S_0} - \beta^0_{S_0}\|_1 \leq \frac{\sqrt{s_0}}{\phi_0} \sqrt{(\hat{\beta} - \beta^0)^\top \hat{\Sigma} (\hat{\beta} - \beta^0)} = \frac{\sqrt{s_0}}{\phi_0} \frac{1}{\sqrt{n}} \|X(\hat{\beta} - \beta^0)\|_2 \tag{64}$$

□

C.2 Static Oracle Inequality

Theorem 3. *(Oracle Inequality of LASSO minimizor)*

On good event $\mathcal{F}$, $\lambda \geq 2\lambda_0$, and compatability condition associate with gram matrix $\hat{\Sigma}$ holds,

$$\frac{1}{n} \|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda \|\hat{\beta} - \beta^0\|_1 \leq \frac{4 s_0}{\phi_0^2} \lambda^2 \tag{65}$$

Proof. Plus both side of basic inequality on good event (48) an addition term $\lambda \|\hat{\beta}_{S_0} - \beta^0_{S_0}\|$ to get

$$\frac{2}{n} \|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda \|\hat{\beta} - \beta^0\|_1 \leq 4\lambda \|\hat{\beta}_{S_0} - \beta^0_{S_0}\|_1 \tag{66}$$

Input lemma 6 to get

$$\frac{2}{n} \|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda \|\hat{\beta} - \beta^0\|_1 \leq 4\lambda \frac{\sqrt{s_0}}{\phi_0} \cdot \frac{1}{\sqrt{n}} \|X(\hat{\beta} - \beta^0)\|_2 \tag{67}$$

Set $u = \frac{1}{\sqrt{n}} \|X(\hat{\beta} - \beta^0)\|_2$ and $v = \lambda \frac{\sqrt{s_0}}{\phi_0}$. Note $(u - 2v)^2 \geq 0$ implies $4uv \leq u^2 + 4v^2$ to get

$$\frac{2}{n} \|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda \|\hat{\beta} - \beta^0\|_1 \leq \frac{\|X(\hat{\beta} - \beta^0)\|_2^2}{n} + 4\lambda^2 \frac{s_0}{\phi_0^2} \tag{68}$$

Reorganize the terms to get

$$\frac{1}{n} \|X(\hat{\beta} - \beta^0)\|_2^2 + \lambda \|\hat{\beta} - \beta^0\|_1 \leq \frac{4 s_0}{\phi_0^2} \lambda^2. \tag{69}$$

□

D Checking Optimal Allocation Condition

Now we show two types of sample set–teamwork sample set and all sample set-produced from our proposed data collection protocol both satisfies the template condition.

The following lemmas are used to prove template condition of teamwork sample set and all sample set(lemma 7 and lemma 10).

D.1 Teamwork Sample Set

Lemma 7. *For any decision epoch $t \geq (Kq)^2$, the teamwork sample set for treatment w_k up to time t, $\mathcal{D}_{[t],w_k}$, is a template sample set of rate p_*, with probability at least $1 - \frac{2}{t^4}$.*

Proof of Lemma 7. To check (i): Lemma 8, $q_0 \geq \frac{6\log d}{Np_* C_2^2(\phi_0)^2}$ and $t > (Kq)^2 > 3$ imply

$$|\mathcal{D}_{[t],k}| \geq \frac{1}{2}Nq\log t \geq 2Nq_0 > \frac{6\log d}{p_* C_2^2(\phi_0)^2}. \tag{70}$$

To check (ii): Lemma 9 shows that, for $t \geq (Kq)^2$, we have

$$P(\frac{|\mathcal{D}^\sharp_{[t],k}|}{|\mathcal{D}_{[t],k}|} \geq \frac{p_*}{2}) \geq 1 - \frac{2}{t^4} \tag{71}$$

Lemma 8. *(Size of Teamwork Sample Set) If $t \geq (Kq)^2$, then*

$$\frac{1}{2}Nq\log t \leq |\mathcal{D}_{[t],k}| \leq 6Nq\log t.$$

Proof. First we note

$$\mathcal{T}_{[t],k} = \mathcal{T}_{\cdot,k} \cap [t] = \cup_{n\geq 0}(\mathcal{T}_{n,k} \cap [t]).$$

At $t \in \mathcal{T}_{n,k}$, we have finished round $0, 1, 2, \cdots, n-1$ teamwork stage for arm k, each of size Nq, therefore

$$nNq \leq |\mathcal{D}_{[t],k}| \leq (n+1)Nq.$$

With this, our task becomes to derive the lower bound and upper bound for n and $n+1$ in terms of $\log t$ by using the condition $t \geq (Kq)^2$

For $t \in \mathcal{T}_{k,n}$,we have

$$(2^n - 1)Kq + 1 \leq t \leq (2^n)Kq,$$

which means

$$\log_2(\frac{t}{Kq}) \leq n \leq \log_2(\frac{t}{Kq} + 1) + 1.$$

Use condition $t \geq (Kq)^2$, one have $\log_2(Kq) \leq \frac{1}{2}\log_2(t)$ and hence

$$n \geq \frac{1}{2}\log_2 t.$$

On the other hand, we have

$$n + 1 \leq \log_2(\frac{t}{Kq} + 1) + 1 \leq \frac{\log(2(t+\sqrt{t}))}{\log 2} \leq 6\log t.$$

□

Lemma 9. *If $t \geq (Kq)^2$, then $P(\frac{|\mathcal{D}^\natural_{[t],k}|}{|\mathcal{D}_{[t],k}|} \geq \frac{p_*}{2}) \geq 1 - \frac{2}{t^4}$.*

Proof. Apply $P(|y-\mu| > \frac{\mu}{2}) < 2\exp[-0.1\mu]$ in Alon and Spencer (2004) to the indicator random variable $I((i,s) \in \mathcal{D}^\natural_{[t],k})$ for all $(i,s) \in \mathcal{D}_{[t],k}$ and using $\mu = E[\sum_{(i,s)\in\mathcal{D}_{[t],k}} I[(i,s) \in \mathcal{D}^\natural_{[t],k}]] \geq p_*|\mathcal{D}_{[t],k},|$ we get $P(|\mathcal{D}^\natural_{[t],k}| < \frac{p_*}{2}|\mathcal{D}_{[t],k}|) < 2e^{-\frac{p_*}{10}|\mathcal{D}_{[t],k}|}$ Therefore, by our control of the size of $|\mathcal{D}_{[t],k}|$ and the choice of q_0, we have $P(|\mathcal{D}^\natural_{[t],k}| < \frac{p_*}{2}|\mathcal{D}_{[t],k}|) < 2e^{-\frac{p_*}{5}q_0\log t} \leq \frac{2}{t^4}$. □

D.2 All Sample Set

We set $\mathcal{S} = \mathcal{T} \cup \mathcal{E}$ in this subsection.

Lemma 10. *For any decision epoch $t \geq C_5$, the all sample set for treatment w_k up to t, $\mathcal{S}_{[t],w_k}$, is a template sample set of rate $\frac{p_*}{2}$, with probability at least* $1 - \exp[-\frac{tp_*^2}{128}]$.

Proof of Lemma 10 To check (i): Lemma 8, $q_0 \geq \frac{6 \log d}{N p_* C_2^2(\phi_0)^2}$ and $t > C_5 > 3$ imply

$$|\mathcal{S}_{[t],k}| \geq |\mathcal{D}_{[t],k}| \geq \frac{12 \log d}{p_* C_2^2(\phi_0)^2} = \frac{6 \log d}{\frac{p_*}{2} C_2^2(\phi_0)^2}. \tag{72}$$

To check (ii): Lemma 11 shows that, for $t \geq C_5$, we have

$$P(\frac{|\mathcal{S}^\sharp_{[t],k}|}{|\mathcal{S}_{[t],k}|} \geq \frac{1}{2}\frac{p_*}{2}) \geq 1 - \exp(-\frac{p_*^2}{128} \cdot t) \tag{73}$$

Lemma 11. *For $t > C_5$,*

$$P(\frac{|\mathcal{S}^\sharp_{[t],k}|}{|\mathcal{S}_{[t],k}|} \geq \frac{1}{2}\frac{p_*}{2}) \geq 1 - \exp(-\frac{p_*^2}{128} \cdot t) \tag{74}$$

Proof. We start from noting the fact that the all sample set for treatment w_k, $\mathcal{S}_{[t],k}$, can have at most t elements up to time t ($|\mathcal{S}_{[t],k}| \leq t$) implies

$$P(\frac{|\mathcal{S}^\sharp_{[t],k}|}{|\mathcal{S}_{[t],k}|} < \frac{1}{2}\frac{p_*}{2}) \geq P(\frac{|\mathcal{S}^\sharp_{[t],k}|}{t} < \frac{p_*}{4}) = P(|\mathcal{S}^\sharp_{[t],k}| < \frac{p_*}{4} \cdot t) \tag{75}$$

To handle RHS, we note that the size of $\mathcal{S}^\sharp_{[t],k}$ admits a representation

$$|\mathcal{S}^\sharp_{[t],k}| = \sum_{s=1}^{t} \sum_{i \in N(s)} I((X_{(i,s)}, Y_{(i,s)}) \in \mathcal{S}^\sharp_{[t],k}). \tag{76}$$

The strategy to utilize such representation is first to construct a martingale difference sequence and then apply Azuma's inequality to attain desired result.

First, all samples been collected in $\mathcal{S}^\sharp_{[t],k}$ are optimal allocation in selfish stage given good event happens. Thus, whether a sample $(X_{i,s}, Y_{i,s})$ belongs to $\mathcal{S}^\sharp_{[t],k}$ has a representation

$$I((X_{i,s}, Y_{i,s}) \in S^\sharp_{[t],k}) = I(E_{s-1}) I(X_{(i,s)} \in U_{w_k}) I(s \notin \mathcal{T}_{[t],\cdot}). \tag{77}$$

Recall that samples in $S^\sharp_{[s],k}$ also satisfies model assumption and hence can be written as $Y = X^\top \beta + \epsilon$. Let $\mathcal{G}_s$ be the sigma algebra generated by the first $N(s) \equiv |S^\sharp_{[s],k}|$ rows of the design matrix X and the first $N(s)$ entries of the noise vector ϵ, and let $\mathcal{G}_0 = \phi$. With this, $I(E_{s-1})$ is $\mathcal{G}_{s-1}$ measurable; $I(X_{(i,s)} \in U_{w_k})$ is $\mathcal{G}_s$ measurable and independent of $\mathcal{G}_{s-1}$; $I(s \notin \mathcal{T}_{[t],\cdot})$ is deterministic by planning of teamwork stage. Follow the Doob's martingale construction, define

$$M_s = E[|\mathcal{S}^\natural_{[t],k}||\mathcal{G}_s] \tag{78}$$

for all $s \in [t] \cup \{0\}$. The resulting sequence $M_0, M_1, \cdots, M_t$ is a martingale adapted to the filtration $\mathcal{G}_0 \subset \mathcal{G}_1 \subset \cdots \mathcal{G}_t$ with $M_0 = E[|\mathcal{S}^\natural_{[t],k}|]$ and $M_t = |\mathcal{S}^\natural_{[t],k}|$. The desired martingale differences is thus $M_s - M_{s-1}$.

Now since the martingale differences $M_s - M_{s-1}$ are bounded by $N(s) - N(s-1)$, the Azuma's inequality, (see. Theorem 7.2.1 from Alon and Spencer 1992), to obtain for all $\eta > 0$,

$$P(|\mathcal{S}^\natural_{[t],k}| < E(|\mathcal{S}^\natural_{[t],k}|) - \eta) \leq \exp(-\frac{\eta^2}{2N(t)}). \tag{79}$$

Now a lower bound for expected size of $\mathcal{S}^{\natural}_{[t],k}$ follows from adopted policy that

$$E[|\mathcal{S}^{\sharp}_{[t],k}|] = \sum_{s=1}^{t}\sum_{i\in N(s)} P((X_{(i,s)},Y_{(i,s)})\in \mathcal{S}^{\sharp}_{[t],k}) \geq [t-|\mathcal{T}_{[t],\cdot}|-(Kq)^2]\frac{p_*}{2}$$

$$\geq [t-6KNq\log t-(Kq)^2]\frac{p_*}{2} \geq \frac{3p_*}{8}t,$$

where the last inequality from the definition of constant C_5. Thus, taking $\eta = \frac{p_*}{8}t$, we have

$$P(|\mathcal{S}^{\sharp}_{[t],k}| < \frac{p_*}{4}t) \leq \exp(-\frac{p_*^2}{128}t). \tag{80}$$

□

E Deviation Inequalities of Teamwork LASSO and ALL LASSO

E.1 Teamwork LASSO-Proof in Corollary 1

Proof. Note $C_1(\frac{\phi_1\sqrt{r}}{2}) = \frac{r^2}{16}C_1(\phi_1)$. Apply Theorem 8 for $\chi = \frac{h}{4x_{\max}}$ and $r = p_*$. First, $q_0 \geq \frac{512x_{\max}^2}{NC_1(\phi_1)p_*^2h^2}$ and lemma 8 imply

$$-C_1(\frac{\phi_1\sqrt{p_*}}{2})|\mathcal{D}_{[t],w_k}|\chi^2+\log d \leq \frac{NC_1(\phi_1)p_*^2h^2}{128x_{\max}^2}\log t\cdot q_0 \leq -4\log t. \tag{81}$$

Second, $|\mathcal{A}^{\sharp}_{w_k}| \geq \frac{r}{2}|\mathcal{A}_{w_k}| \geq \frac{p_*}{2}Nq_0$ and $q_0 \geq \frac{8}{NC_2(\phi_1)^2p_*}$ and it implies

$$-|\mathcal{D}^{\sharp}_{[t],w_k}|C_2(\phi_1)^2 \leq -\frac{NC_2(\phi_1)^2p_*}{2}\cdot q_0 \leq -4\log t \tag{82}$$

Last, we find

$$P(\|\hat{\beta}_{w_k}(\mathcal{D}_{[t],k},\lambda_1)-\beta_{w_k}\|_1 > \frac{h}{4x_{\max}}) \tag{83}$$

$$\leq \quad P(\|\hat{\beta}_{w_k}(\mathcal{D}_{[t],k},\lambda_1)-\beta_{w_k}\|_1 > \frac{h}{4x_{\max}}, \frac{|\mathcal{D}^{\sharp}_{[t],k}|}{|\mathcal{D}_{[t],k}|} \geq \frac{p_*}{2}) + P(\frac{|\mathcal{D}^{\sharp}_{[t],k}|}{|\mathcal{D}_{[t],k}|} < \frac{p_*}{2}) \tag{84}$$

$$\leq \quad 2\cdot\frac{1}{t^4}+\frac{1}{t^4}+\frac{2}{t^4} = \frac{5}{t^4} \tag{85}$$

E.2 All LASSO–Proof in Corollary 2

Proof. Note $C_1(\frac{\phi_1\sqrt{r}}{2}) = \frac{r^2}{16}C_1(\phi_1)$. Apply Theorem 1 for $\chi = \frac{16}{\sqrt{p_*^3C_1(\phi_0)}}\sqrt{\frac{\log t+\log d}{t}}$ and $r = \frac{p_*}{2}$. First, $|\mathcal{S}_{[t],w_k}| \geq \frac{p_*t}{4}$ and lemma 8 imply

$$-C_1(\frac{\phi_1\sqrt{p_*/2}}{2})|\mathcal{S}_{[t],w_k}|\chi^2+\log d \leq -\frac{p_*^2}{64}C_1\cdot\frac{p_*t}{4}\cdot 256\frac{\log t+\log d}{tp_*^3C_1}+\log d = -\log t. \tag{86}$$

Second, $|\mathcal{S}^{\sharp}_{[t],w_k}| \geq \frac{p_*}{4}|\mathcal{S}_{[t],w_k}| \geq \frac{p_*^2t}{16}$ and $C_2^2 \geq \frac{1}{2}$ imply

$$-|\mathcal{S}^{\sharp}_{[t],w_k}|C_2(\phi_1)^2 \leq -\frac{p_*^2}{32}\cdot t \tag{87}$$

Last, we find

$$P(\|\hat{\beta}_{w_k}(\mathcal{D}_{[t],k},\lambda_1)-\beta_{w_k}\|_1 > \frac{16}{\sqrt{p_*^3C_1(\phi_0)}}\sqrt{\frac{\log t+\log d}{t}}) \tag{88}$$

$$\leq \quad P(\|\hat{\beta}_{w_k}(\mathcal{D}_{[t],k},\lambda_1)-\beta_{w_k}\|_1 > \frac{16\sqrt{\frac{\log t+\log d}{t}}}{\sqrt{p_*^3C_1(\phi_0)}}, \frac{|\mathcal{S}^{\sharp}_{[t],k}|}{|\mathcal{S}_{[t],k}|} \geq \frac{p_*}{2}) + P(\frac{|\mathcal{S}^{\sharp}_{[t],k}|}{|\mathcal{S}_{[t],k}|} < \frac{p_*}{2}) \tag{89}$$

$$\leq \quad 2\cdot\frac{1}{t}+\exp(-\frac{p_*^2}{32}\cdot t)+\exp(-\frac{p_*^2}{128}\cdot t) \leq 2(\frac{1}{t}+\exp(-\frac{p_*^2}{32}\cdot t)) \tag{90}$$

F Regret Analysis

We show the properties of $\hat{K}(x)$ for $x \in \mathcal{X}$ and for $x \in U_w$ of a available treatment w. In words, for any given observed covariate $x \in \mathcal{X}$, Teamwork LASSO excludes those sub-optimal treatment of x up to tolerance level h. If $x \in U_w$, then Teamwork excludes all treatment other than the optimal treatment of x. Therefore, the probability of random covariate X belongs to U_w matters.

F.1 Proof of lemma 12

Lemma 12. *(For $x \in \mathcal{X}$) Suppose the $(t-1)$th decision epoch is in selfish stage and event E_{t-1} holds. Then for each available treatment $w_i \in \mathcal{W}$ and any possible observed covariate $x \in \mathcal{X}$, the estimated optimal treatment candidate set contains the optimal treatment of x: $w^*(x) \equiv \arg\max_{w\in\mathcal{W}} \langle x, \beta_w \rangle$ and no any sub-optimal treatment $w \in \mathcal{W}_{sub}$. That is,*

$$w^*(x) \in \widehat{K}(x) \quad \text{and} \quad w \notin \widehat{K}(x) \;\; \text{for all} \;\; w \in \mathcal{W}_{sub} \tag{91}$$

Proof. First, we show $w^*(x) \in \widehat{K}(x)$. Note at the tth decision epoch, the optimal treatment suggested by Teamwork LASSO is $w^{\text{team}}(x) \equiv \arg\max_{w\in\mathcal{W}} x^\top \widehat{\beta}_w(\mathcal{D}_{[t-1],w}, \lambda_1)$. Since E_{t-1} holds, it implies $x^\top \widehat{\beta}_w(\mathcal{D}_{[t-1],w}, \lambda_1) - x^\top \beta_w < x_{\max} \cdot \frac{h}{4x_{\max}} = \frac{h}{4}$ for all available treatment w, which includes w^* and w^{team}.

$$\begin{aligned}
& x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{team}}}) - x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^*}) && (92)\\
= \;& (x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{team}}}) - x^\top \beta_{w^{\text{team}}}) + (x^\top \beta_{w^{\text{team}}} - x^\top \beta_{w^*}) + (x^\top \beta_{w^*} - x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^*}) && (93)\\
\leq \;& \frac{h}{4} + 0 + \frac{h}{4} = \frac{h}{2}, && (94)
\end{aligned}$$

where the last inequality is from the definition of $w^*(x)$ that $x^\top \beta_{w^*} - x^\top \beta_{w^{\text{team}}} < 0$.

Second, we show $w_{\text{sub}} \notin \widehat{K}(x)$ for all $w_{\text{sub}} \in \mathcal{W}_{\text{sub}}$. Since E_{t-1} holds, it implies $x^\top \widehat{\beta}_w(\mathcal{D}_{[t-1],w}, \lambda_1) - x^\top \beta_w > -x_{\max} \cdot \frac{h}{4x_{\max}} = -\frac{h}{4}$ for all available treatment w, which includes w^{team} and w^{sub}.

$$\begin{aligned}
& x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{team}}}) - x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{sub}}}) && (95)\\
\geq \;& x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^*}) - x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{sub}}}) && (96)\\
= \;& (x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^*}) - x^\top \beta_{w^*}) + (x^\top \beta_{w^*} - x^\top \beta_{w^{\text{sub}}}) + (x^\top \beta_{w^{\text{sub}}} - x^\top \widehat{\beta}(\mathcal{D}_{[t-1],w^{\text{sub}}}) && (97)\\
\geq \;& -\frac{h}{4} + h + -\frac{h}{4} = \frac{h}{2}, && (98)
\end{aligned}$$

where the last inequality is from the definition of $\mathcal{W}_{\text{sub}}$ that $x^\top \beta_{w^*} - x^\top \beta_{w^{\text{sub}}} > h$. □

F.2 Proof of lemma 13

Lemma 13. *(For $x \in U_{w_i}$) Suppose the $(t-1)$th decision epoch is in selfish stage and event E_{t-1} holds. Then for each available treatment $w_i \in \mathcal{W}$, if a observed covariate x belongs to U_{w_i}, then the estimated optimal treatment candidate set contains only treatment w_i, that is*

$$\widehat{K}(x) = \{w_i\}. \tag{99}$$

Proof. For every treatment w_j other than w_i, since $x \in U_{w_i}$, definition of U_{w_i} implies $x^\top \beta_{w_i} - x^\top \beta_{w_j} > h$; since E_{t-1} holds, it implies $x^\top \widehat{\beta}_w(\mathcal{D}_{[t-1],w}, \lambda_1) - x^\top \beta_w > -x_{\max} \cdot \frac{h}{4x_{\max}} = -\frac{h}{4}$. Combine them to obtain, for every treatment w_j other than w_i

$$\begin{aligned}
& x^\top \widehat{\beta}_{w_i}(\mathcal{D}_{[t-1],w_i}, \lambda_1) - x^\top \widehat{\beta}_{w_j}(\mathcal{D}_{[t-1],w_j}, \lambda_1) && (100)\\
= \;& x^\top [\widehat{\beta}_{w_i}(\mathcal{D}_{[t-1],w_i}, \lambda_1) - \beta_{w_i}] - x^\top [\widehat{\beta}_{w_j}(\mathcal{D}_{[t-1],w_j}, \lambda_1) - \beta_{w_j}] + x^\top [\beta_{w_i} - \beta_{w_j}] && (101)\\
\geq \;& -\frac{h}{4} - \frac{h}{4} + h = \frac{h}{2}. && (102)
\end{aligned}$$

That is, for every treatment w_j other than w_i,

$$x^\top \widehat{\beta}_{w_i}(\mathcal{D}_{[t-1],w_i}, \lambda_1) \geq x^\top \widehat{\beta}_{w_j}(\mathcal{D}_{[t-1],w_j}, \lambda_1) + \frac{h}{2}. \tag{103}$$

Therefore, by construction of optimal treatment candidate set, we conclude $\widehat{K}(x) = \{w_i\}$. □

F.3 Regret bound for case (4)

Lemma 14.

$$f(t) = [4Kbx_{\max} + C_3(\phi_0, p_*) \cdot \log d]\frac{1}{t} + 8Kbx_{\max}\exp[-\frac{p_*^2 C_2(\phi_0)^2}{32} \cdot t] + C_3(\phi_0, p_*)\frac{\log t}{t} \tag{104}$$

Proof. Without loss of generality, for a observed covariate vector $x_{i,t}$ of the ith user at the tth decision epoch, assume w_1 is the optimal treatment, that is $x_{i,t}^\top \beta_{w_1} = \max_{w\in\mathcal{W}} x_{i,t}^\top \beta_w$. First, we note that the instantaneous regret occurs if we allocate treatment other than w_1 to covariate x. This happens when $x^\top \widehat{\beta}(\mathcal{S}_{[t-1],w}) > x^\top \widehat{\beta}(\mathcal{S}_{[t-1],w_1})$ for some treatments w. This observation suggests

$$\begin{aligned} r_{i,t} &= E[\sum_{w_k\in\hat{K}(x_{i,t})} x_{i,t}^\top(\beta_{w_1} - \beta_{w_k}) I(\pi(x_{i,t} = w_k)] && (105)\\ &\leq E[\sum_{w_k\in\hat{K}(x_{i,t})} x_{i,t}^\top(\beta_{w_1} - \beta_{w_k}) I(x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_k}) > x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_1})] && (106)\end{aligned}$$

Second, to handle RHS, define a function $g(x) \equiv x^\top(\beta_{w_1} - \beta_{w_k})$ consider the set

$$B_{w_k} \equiv \{x | x^\top(\beta_{w_1} - \beta_{w_k}) > 2\delta x_{\max}\}. \tag{107}$$

The boundness assumption on observed covariate x and efficacy parameter β_w suggests $g(x) \leq 2bx_{\max}$ for all $x \in B_{w_k}$; the definition of set B_{w_k} suggests $g(x) \leq 2\delta x_{\max}$ for all $x \in B_{w_k}^c$. This observation suggests

$$\begin{aligned} r_{i,t} &\leq |\hat{K}(x_{i,t})| \cdot 2bx_{\max} \cdot E[I(x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_k}) > x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_1}) I(x_{i,t} \in B_{w_k})] && (108)\\ &+ |\hat{K}(x_{i,t})| \cdot 2\delta x_{\max} \cdot E[I(x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_k}) > x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_1}) I(x_{i,t} \in B_{w_k}^c)] && (109)\\ &\leq K2bx_{\max} E[I(x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_k}) > x_{i,t}^\top \widehat{\beta}(\mathcal{S}_{[t],w_1}) I(x_{i,t}^\top(\beta_{w_1} - \beta_{w_k}) > 2\delta x_{\max})] && (110)\\ &+ K2\delta x_{\max} E[I(x_{i,t}^\top(\beta_{w_1} - \beta_{w_k}) \leq 2\delta x_{\max})] && (111)\end{aligned}$$

Third, we handle equation (110) and (111). We note the marginal condition implies

$$(111) = K2\delta x_{\max} P(X^\top(\beta_{w_1} - \beta_{w_k}) \leq 2\delta x_{\max}) \leq C_0 \cdot 2\delta x_{\max}. \tag{112}$$

Based on this observation, we have

$$\begin{aligned} (110) &\leq K2bx_{\max} \cdot (P(\|\beta_{w_1} - \widehat{\beta}_{w_1}(\mathcal{S}_{[t],w_1})\|_1 > \delta) + P(\|\widehat{\beta}_{w_k}(\mathcal{S}_{[t],w_k}) - \beta_{w_k}\|_1 > \delta)) && (113)\\ &\leq K2bx_{\max} \cdot 2 \cdot (\frac{1}{t} + 2\exp(-\frac{p_*^2 C_2(\phi_0)^2}{32} \cdot t)) && (114)\end{aligned}$$

Last, combine above results and take $\delta = 16\sqrt{\frac{\log t + \log d}{p_*^3 C_1 t}}$, we have

$$\begin{aligned} & r_{i,t} && (115)\\ \leq\ & K2bx_{\max} \cdot 2 \cdot (\frac{1}{t} + 2\exp(-\frac{p_*^2 C_2(\phi_0)^2}{32} \cdot t)) + 2\delta x_{\max} \cdot C_0 \cdot 2\delta x_{\max}) && (116)\\ =\ & K4bx_{\max}(\frac{1}{t} + 2\exp(-\frac{p_*^2 C_2(\phi_0)^2}{32} \cdot t)) + 4\delta^2 x_{\max}^2 \cdot C_0) && (117)\\ =\ & [4Kbx_{\max} + C_3(\phi_0, p_*)\log d]\frac{1}{t} + 8Kbx_{\max}\exp[-\frac{p_*^2 C_2(\phi_0)^2}{32} \cdot t] + C_3(\phi_0, p_*)\frac{\log t}{t}, && (118)\end{aligned}$$

as desired. □

F.4 Full Regret Bound–Proof of Theorem 2

The regret can be bounded by:

$$R_T = \sum_{t\in[T]}\sum_{i\in[N]} r_{i,t} \tag{119}$$

$$= \sum_{t\in[C_5]}\sum_{i\in[N]} r_{i,t} + \sum_{t\in[C_5:T]\cap\mathcal{T}}\sum_{i\in[N]} r_{i,t} + \sum_{t\in[C_5:T]\cap\mathcal{T}^\lrcorner}\sum_{i\in[N]} r_{i,t} \tag{120}$$

$$\leq N\cdot C_5\cdot 2bx_{\max} + N\cdot|\mathcal{T}|\cdot 2bx_{\max} + N\cdot\sum_{t\in[C_5:T]\cap\mathcal{T}^\lrcorner}[\frac{K}{t^4}\cdot 2bx_{\max} + f(t)] \tag{121}$$

$$\leq NC_5 2bx_{\max} + N(6q\log TK)2bx_{\max} + NK2bx_{\max}\int_1^T \frac{1}{t^4}dt + N\cdot\int_1^T f(t)dt \tag{122}$$

$$\leq N\cdot\{2bx_{\max}\cdot[C_5 + 6qK\log T + K] \tag{123}$$

$$+ [4Kbx_{\max} + C_3(\phi_0,p_*)\cdot\log d]\log T + 8Kbx_{\max}C_4 + C_3(\phi_0,p_*)(\log T)^2\} \tag{124}$$

G Constants

Here we list the constants that appear in the proof.

- $C_1(\phi_0) \equiv \frac{\phi_0^4}{512 s_0^2\sigma^2 x_{\max}^2}$
- $C_2 \equiv \min\{\frac{1}{2}, \frac{\phi_0^2}{256 s_0 x_{\max}^2}\}$
- $C_3 \equiv \frac{1024 K C_0 x_{\max}^2}{p_*^3 C_1}$
- $C_4 \equiv \frac{8Kbx_{\max}}{1-\exp(-\frac{p_*^2}{32})}$
- $C_5 \equiv \{t\in\mathbb{Z}^+ | t \geq 24Kq\log t + 4(Kq)^2\}$
- $q_0 \equiv \max\{\frac{20}{Np_*}, \frac{4}{Np_* C_2^2}, \frac{3\log d}{Np_* C_2^2}, \frac{1024 x_{\max}^2 \log d}{Nh^2 p_*^2 C_1}\}$.

H Experiment

In Figure. 3, we compare our `Teamwork LASSO Bandit` with batch size $N = 4$ and $N = 12$ to the `LASSO Bandit` in Bastani and Bayati (2020). In the attached plot, covariate dimension d = 200, 500 and 1000, number of treatments (arms) K =3, the length of exploration phase q = 1,2,3,4,5,6 with a total number of decisions 5000. N is the batch size, where N=1 corresponds to LASSO Bandit and N=4, 12 corresponds our Teamwork LASSO Bandit. We run 100 replications for each setting.

Remark on cumulative regret and covariate vector dimension. In the experiment, we increase the covariate vector dimension from 200, 500 to 1000. The performance of high update frequency algorithm is more sensitive to the increase in covariate dimension than our low update frequency algorithm.

Remark on the length of exploration phase q. In real world practice, the length of exploration phase q is pre-specified and then an explore-exploitation policy follows the choice of q. Given the same total number of decisions, it is often the case that one prefers a smaller value of q, which means fewer regret from exploration and is more time efficient in the sense that more rounds of explore-exploit can be done.

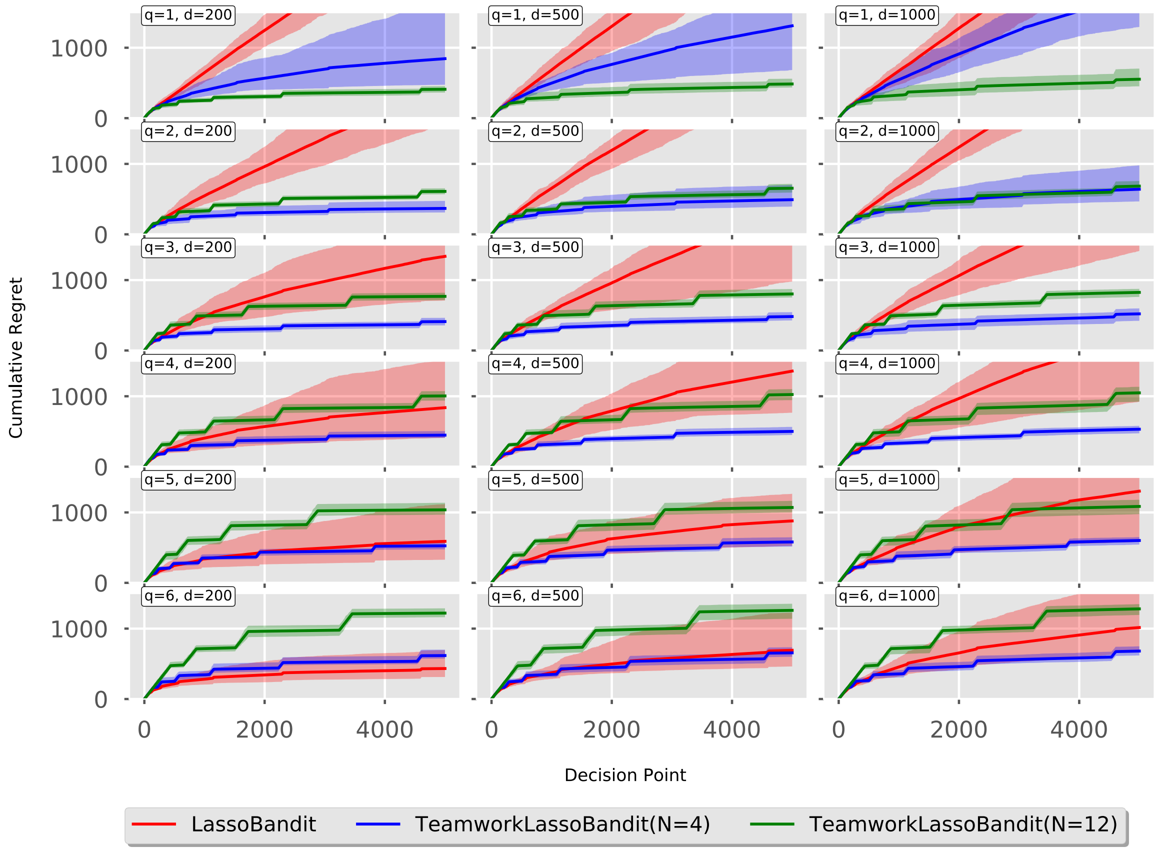

Figure 3: Comparison of our `Teamwork LASSO Bandit` with batch size $N = 4$ and $N = 12$ to the `LASSO Bandit` in Bastani and Bayati (2020). The error bars represent the maximum and minimum of the regret among 100 replications.