Supplementary Material:
Stay With Me: Lifetime Maximization Through Heteroscedastic Linear Bandits With Reneging

A. Appendix

A.1. Proof of Lemma 2

Proof. Recall that $\boldsymbol{V}_n = \left(\boldsymbol{X}_n^\top \boldsymbol{X}_n + \lambda \boldsymbol{I}_d\right)$. Note that

$$\widehat{\phi}_n = (\boldsymbol{X}_n^\top \boldsymbol{X}_n + \lambda I_d)^{-1} \boldsymbol{X}_n^\top f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) \quad (39)$$
$$= \boldsymbol{V}_n^{-1} \boldsymbol{X}_n^\top f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) \quad (40)$$
$$= \boldsymbol{V}_n^{-1} \boldsymbol{X}_n^\top \left(f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) - \boldsymbol{X}_n \phi_* + \boldsymbol{X}_n \phi_*\right) \quad (41)$$
$$+ \lambda \boldsymbol{V}_n^{-1} \phi_* - \lambda \boldsymbol{V}_n^{-1} \phi_* \quad (42)$$
$$= \boldsymbol{V}_n^{-1} \boldsymbol{X}_n^\top \left(f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) - \boldsymbol{X}_n \phi_*\right) - \lambda \boldsymbol{V}_n^{-1} \phi_* + \phi_*. \quad (43)$$

Therefore, for any $x \in \mathbb{R}^d$, we know

$$|x^\top \widehat{\phi}_n - x^\top \widehat{\phi}_*| \quad (44)$$
$$= |x^\top \boldsymbol{V}_n^{-1} \boldsymbol{X}_n^\top \left(f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) - \boldsymbol{X}_n \phi_*\right) - \lambda x^\top \boldsymbol{V}_n^{-1} \phi_*| \quad (45)$$
$$\leq \|x\|_{\boldsymbol{V}_n{}^{-1}} \Big(\lambda \|\phi_*\|_{\boldsymbol{V}_n^{-1}} \quad (46)$$
$$+ \left\|\boldsymbol{X}_n^\top \left(f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) - \boldsymbol{X}_n \phi_*\right)\right\|_{\boldsymbol{V}_n{}^{-1}}\Big). \quad (47)$$

Moreover, by rewriting $\widehat{\varepsilon} = \widehat{\varepsilon} - \varepsilon + \varepsilon$, we have

$$f^{-1}(\widehat{\varepsilon} \circ \widehat{\varepsilon}) \quad (48)$$
$$= f^{-1}\big((\widehat{\varepsilon} - \varepsilon + \varepsilon) \circ (\widehat{\varepsilon} - \varepsilon + \varepsilon)\big) \quad (49)$$
$$= f^{-1}(\varepsilon \circ \varepsilon) + M_f^{-1}\Big(2\big(\varepsilon \circ \boldsymbol{X}_n(\theta_* - \widehat{\theta}_n)\big) \quad (50)$$
$$+ \big(\boldsymbol{X}_n(\theta_* - \widehat{\theta}_n) \circ \boldsymbol{X}_n(\theta_* - \widehat{\theta}_n)\big)\Big), \quad (51)$$

where (50)-(51) follow from the fact that both $f(\cdot)$ and $f^{-1}(\cdot)$ are linear with a slope M_f and M_f^{-1}, respectively, as described in Section 3. Therefore, by (44)-(51) and the Cauchy-Schwarz inequality, we have

$$|x^\top \widehat{\phi}_n - x^\top \widehat{\phi}_*| \leq \|x\|_{\boldsymbol{V}_n{}^{-1}} \Big\{\lambda \|\phi_*\|_{\boldsymbol{V}_n^{-1}} \quad (52)$$
$$+ \left\|\boldsymbol{X}_n^\top \left(f^{-1}(\varepsilon \circ \varepsilon) - \boldsymbol{X}_n \phi_*\right)\right\|_{\boldsymbol{V}_n{}^{-1}} \quad (53)$$
$$+ 2M_f^{-1} \left\|\boldsymbol{X}_n^\top \left(\varepsilon \circ \boldsymbol{X}_n(\theta_* - \widehat{\theta}_n)\right)\right\|_{\boldsymbol{V}_n{}^{-1}} \quad (54)$$
$$+ M_f^{-1} \left\|\boldsymbol{X}_n^\top \left(\boldsymbol{X}_n(\theta_* - \widehat{\theta}_n) \circ \boldsymbol{X}_n(\theta_* - \widehat{\theta}_n)\right)\right\|_{\boldsymbol{V}_n{}^{-1}} \Big\}. \quad (55)$$

□

A.2. Proof of Lemma 3

We first introduce the following useful lemmas.

Lemma A.1 (Lemma 8.2 in (Erdős et al., 2012)) *Let $\{a_i\}_{i=1}^N$ be N independent random complex variables with zero mean and variance σ^2 and having uniform sub-exponential decay, i.e., there exists $\kappa_1, \kappa_2 > 0$ such that*

$$\mathbb{P}\{|a_i| \geq x^{\kappa_1}\} \leq \kappa_2 e^{-x}. \quad (56)$$

We use a^{H} to denote the conjugate transpose of a. Let $a = (a_1, \cdots, a_N)^\top$, let $\overline{a_i}$ denote the complex conjugate of a_i, for all i, and let $\boldsymbol{B} = (B_{ij})$ be a complex $N \times N$ matrix. Then, we have

$$\mathbb{P}\Big\{|a^{\mathrm{H}} \boldsymbol{B} a - \sigma^2 \mathrm{tr}(\boldsymbol{B})| \geq s\sigma^2 \Big(\sum_{i=1}^N |B_{ii}|^2\Big)^{-1/2}\Big\} \quad (57)$$
$$\leq C_1 \exp\Big(-C_2 \cdot s^{1/(1+\kappa_1)}\Big), \quad (58)$$

where C_1 and C_2 are positive constants that depend only on κ_1, κ_2. Moreover, for the standard χ_1^2-distribution, $\kappa_1 = 1$ and $\kappa_2 = 2$.

For any $p \times q$ matrix $\boldsymbol{A}$, we define the induced matrix norm as $\|\boldsymbol{A}\|_2 := \max_{v \in \mathbb{R}^q, \|v\|_2 = 1} \|\boldsymbol{A}v\|_2$.

Lemma A.2

$$\left\|\boldsymbol{V}_n{}^{-1/2} \boldsymbol{X}^\top\right\|_2 \leq 1, \forall n \in \mathbb{N}. \quad (59)$$

Proof. By the definition of induced matrix norm,

$$\left\|\boldsymbol{V}_n{}^{-1/2} \boldsymbol{X}^\top\right\|_2 = \max_{\|v\|_2 = 1} \sqrt{v^\top \boldsymbol{X} \boldsymbol{V}_n{}^{-1} \boldsymbol{X}^\top v} \quad (60)$$
$$= \lambda_{\max}\Big(\boldsymbol{X} \boldsymbol{V}_n{}^{-1} \boldsymbol{X}^T\Big) \quad (61)$$
$$= \lambda_{\max}\Big(\boldsymbol{X}\left(\boldsymbol{X}^T \boldsymbol{X} + \lambda \boldsymbol{I}_d\right)^{-1} \boldsymbol{X}^T\Big) \quad (62)$$
$$\leq \frac{\lambda_{\max}(\boldsymbol{X}^\top \boldsymbol{X})}{\lambda_{\max}(\boldsymbol{X}^\top \boldsymbol{X}) + \lambda} \leq 1, \quad (63)$$

where (63) follows from the singular value decomposition and $\lambda_{\max}(\boldsymbol{X}^\top \boldsymbol{X}) \geq 0$. □

To simplify notation, we use $\boldsymbol{X}$ and $\boldsymbol{V}$ as a shorthand for $\boldsymbol{X}_n$ and $\boldsymbol{V}_n$, respectively. For convenience, we rewrite $\boldsymbol{V}^{-1/2}X^\top = [v_1 \cdots v_n]$ as the matrix of n column vectors $\{v_i\}_{i=1}^n$ (each $v_i \in \mathbb{R}^d$) and show the following property.

Lemma A.3 *Let $v_i \in \mathbb{R}^d$ be the i-th column of the matrix $\boldsymbol{V}^{-1/2}X^\top$, for all $1 \leq i \leq n$. Then, we have*

$$\sum_{i=1}^{n} \|v_i\|_2^2 \leq d. \tag{64}$$

Proof of Lemma A.3. Recall that $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a square matrix. We know

$$\sum_{i=1}^{n} \|v_i\|_2^2 = \text{tr}\Big(\big(\boldsymbol{X}\boldsymbol{V}^{-1/2}\big)\big(\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\big)\Big) \tag{65}$$

$$= \text{tr}\Big(\big(\boldsymbol{V}^{-1/2}\boldsymbol{X}\big)\big(\boldsymbol{X}^\top\boldsymbol{V}^{-1/2}\big)\Big) \tag{66}$$

$$\leq d \cdot \lambda_{\max}\Big(\big(\boldsymbol{V}^{-1/2}\boldsymbol{X}\big)\big(\boldsymbol{X}^\top\boldsymbol{V}^{-1/2}\big)\Big), \tag{67}$$

where (66) follows from the trace of a product being commutative, and (67) follows since the trace is the sum of all eigenvalues. Moreover, we have

$$\lambda_{\max}\big(\big(\boldsymbol{X}\boldsymbol{V}^{1/2}\big)\big(\boldsymbol{X}^\top\boldsymbol{V}^{-1/2}\big)\big) \tag{68}$$

$$= \Big\|\big(\boldsymbol{X}\boldsymbol{V}^{1/2}\big)\big(\boldsymbol{X}^\top\boldsymbol{V}^{-1/2}\big)\Big\|_2 \tag{69}$$

$$\leq \Big\|\big(\boldsymbol{X}\boldsymbol{V}^{1/2}\big)\Big\|_2 \Big\|\big(\boldsymbol{X}^\top\boldsymbol{V}^{-1/2}\big)\Big\|_2 \leq 1, \tag{70}$$

where (70) follows from the fact that the ℓ_2-norm is sub-multiplicative. Therefore, by (65)-(70), we conclude that $\sum_{i=1}^n \|v_i\|_2^2 \leq d$. □

We are now ready to prove Lemma 3.

Proof of Lemma 3. To simplify notation, we use $\boldsymbol{X}$ and $\boldsymbol{V}$ as a shorthand for $\boldsymbol{X}_n$ and $\boldsymbol{V}_n$, respectively. To begin with, we know $f^{-1}(\varepsilon \circ \varepsilon) - \boldsymbol{X}\phi_* = \frac{1}{M_f}((\varepsilon \circ \varepsilon) - f(\boldsymbol{X}\phi_*))$. Therefore, we have

$$\big\|\boldsymbol{X}(f^{-1}(\varepsilon \circ \varepsilon) - \boldsymbol{X}\phi_*)\big\|_{\boldsymbol{V}^{-1}} \tag{71}$$

$$= \frac{1}{M_f}\sqrt{\big(\varepsilon \circ \varepsilon - f(\boldsymbol{X}\phi_*)\big)^\top \boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X}^\top\big(\varepsilon \circ \varepsilon - f(\boldsymbol{X}\phi_*)\big)}, \tag{72}$$

where each element in the vector $(\varepsilon \circ \varepsilon - f(\boldsymbol{X}\phi_*))$ is a centered χ_1^2-distribution with a scaling of $f(\phi_*^\top x_i)$. Defining $\boldsymbol{W} = \text{diag}\big(f(x_1^\top\phi_*), ..., f(x_n^\top\phi_*)\big)$, we have

$$\big\|\boldsymbol{X}(f^{-1}(\varepsilon \circ \varepsilon) - \boldsymbol{X}\phi_*)\big\|_{\boldsymbol{V}^{-1}} \tag{73}$$

$$= \frac{1}{M_f}\Big[\underbrace{\Big(\varepsilon \circ \varepsilon - f(\boldsymbol{X}\phi_*)\Big)^\top \boldsymbol{W}^{-1}}_{\text{mean=0, variance= 2}}\Big(\boldsymbol{W}\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X}^\top\boldsymbol{W}\Big) \tag{74}$$

$$\underbrace{\boldsymbol{W}^{-1}\Big(\varepsilon \circ \varepsilon - f(\boldsymbol{X}\phi_*)\Big)}_{\text{mean=0, variance=2}}\Big]^{1/2}. \tag{75}$$

We use $\eta = \boldsymbol{W}^{-1}\big(\varepsilon\circ\varepsilon - f(\boldsymbol{X}\phi_*)\big)$ as a shorthand and define $\boldsymbol{U} = \big(U_{ij}\big) = \boldsymbol{W}\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X}^\top\boldsymbol{W}$. By Lemma A.1 and the fact that $\varepsilon(x_1), \cdots, \varepsilon(x_n)$ are mutually independent given the contexts $\{x_i\}_{i=1}^n$, we have

$$\mathbb{P}\Big\{|\eta^\top\boldsymbol{U}\eta - 2\cdot\text{tr}(\boldsymbol{U})| \geq 2s\Big(\sum_{i=1}^{n}|\boldsymbol{U}_{ii}|^2\Big)^{1/2}\Big\} \tag{76}$$

$$\leq C_1 \text{exp}(-C_2\sqrt{s}). \tag{77}$$

Recall that $\boldsymbol{V}^{-1/2}X^\top = [v_1 \cdots v_n]$. The trace of $\boldsymbol{U}$ can be upper bounded as

$$\text{tr}(\boldsymbol{U}) = \text{tr}(\boldsymbol{W}\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X}^\top\boldsymbol{W}) \tag{78}$$

$$= \text{tr}\Big(\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{W}\boldsymbol{W}\boldsymbol{X}\boldsymbol{V}^{-1/2}\Big) \tag{79}$$

$$= \sum_{i=1}^{n} f(x_i^\top\phi_*)^2 \cdot \|v_i\|_2^2 \tag{80}$$

$$\leq (\sigma_{\max}^2)^2 \sum_{i=1}^{n} \|v_i\|_2^2 \leq (\sigma_{\max}^2)^2 d, \tag{81}$$

where the last inequality in (81) follows directly from Lemma A.3. Also by the commutative property of the trace operation, we have

$$\sum_{i=1}^{n}|\boldsymbol{U}_{ii}|^2 \overset{(a)}{\leq} \Big(\sum_{i=1}^{n}\boldsymbol{U}_{ii}\Big)^2 \overset{(b)}{\leq} \big((\sigma_{\max}^2)^2 d\big)^2, \tag{82}$$

where (a) follows from $\boldsymbol{U}$ being positive semi-definite (all diagonal elements are nonnegative), and (b) follows from (81). Therefore, by (76)-(82), we have

$$\mathbb{P}\Big\{\eta^\top\boldsymbol{U}\eta \geq 2s\cdot(\sigma_{\max}^2)^2 d + 2(\sigma_{\max}^2)^2 d\Big\} \tag{83}$$

$$\leq C_1 \cdot \text{exp}(-C_2\sqrt{s}). \tag{84}$$

By choosing $s = \Big(\frac{1}{C_2}\ln\frac{C_1}{\delta}\Big)^2$, we have

$$\mathbb{P}\Big\{\eta^\top\boldsymbol{U}\eta \geq 2(\sigma_{\max}^2)^2 d\Big(\Big(\frac{1}{C_2}\ln\frac{C_1}{\delta}\Big)^2 + 1\Big)\Big\} \leq \delta. \tag{85}$$

Therefore, we conclude that with probability at least $1-\delta$, the following inequality holds

$$\left\|\boldsymbol{X}(f^{-1}(\varepsilon\circ\varepsilon)-\boldsymbol{X}\phi_*)\right\|_{\boldsymbol{V}^{-1}} \tag{86}$$

$$\le \frac{1}{M_f}\sqrt{2(\sigma^2_{\max})^2\cdot d\Big(\Big(\frac{1}{C_2}\ln\frac{C_1}{\delta}\Big)^2+1\Big)}. \tag{87}$$

□

A.3. Proof of Lemma 4

We first introduce a useful lemma.

Lemma A.4 (Theorem 4.1 in (Tropp, 2012)) *Consider a finite sequence $\{\boldsymbol{A}_k\}$ of fixed self-adjoint matrices of dimension $d\times d$, and let $\{\gamma_k\}$ be a finite sequence of independent standard normal variables. Let $\sigma^2=\left\|\sum_k \boldsymbol{A}_k^2\right\|_2$. Then, for all $s\ge 0$,*

$$\mathbb{P}\Big\{\lambda_{\max}\Big(\sum_k \gamma_k \boldsymbol{A}_k\Big)\ge s\Big\}\le d\cdot\exp(-\frac{s^2}{2\sigma^2}), \tag{88}$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a square matrix.

Now we are ready to prove Lemma 4.

Proof of Lemma 4. To simplify notation, we use $\boldsymbol{X}$ and $\boldsymbol{V}$ as a shorthand for $\boldsymbol{X}_n$ and $\boldsymbol{V}_n$, respectively. Recall that $\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top=[v_1,v_2,...,v_n]$ and define $\boldsymbol{A}_i=v_iv_i^\top$, for all $i=1,...,n$. Note that $\boldsymbol{A}_i$ is symmetric, for all i. Define an $n\times n$ diagonal matrix $\boldsymbol{D}=\text{diag}(\varepsilon_1,\varepsilon_2,...,\varepsilon_n)$. Then we have:

$$\left\|\boldsymbol{X}^\top\Big(\varepsilon\circ\big(\boldsymbol{X}(\theta_*-\widehat{\theta})\big)\Big)\right\|_{\boldsymbol{V}^{-1}} \tag{89}$$

$$=\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\Big(\varepsilon\circ\big(\boldsymbol{X}(\theta_*-\widehat{\theta})\big)\Big)\right\|_2 \tag{90}$$

$$=\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{D}\boldsymbol{X}(\theta_*-\widehat{\theta})\right\|_2 \tag{91}$$

$$=\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{D}\boldsymbol{X}\boldsymbol{V}^{-1/2}\boldsymbol{V}^{1/2}(\theta_*-\widehat{\theta})\right\|_2 \tag{92}$$

$$\le\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{D}\boldsymbol{X}\boldsymbol{V}^{-1/2}\right\|_2\cdot\left\|\boldsymbol{V}^{1/2}(\theta_*-\widehat{\theta})\right\|_2 \tag{93}$$

$$=\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{D}\boldsymbol{X}\boldsymbol{V}^{-1/2}\right\|_2\cdot\left\|\theta_*-\widehat{\theta}\right\|_{\boldsymbol{V}}. \tag{94}$$

Next, the first term in (94) can be expanded into

$$\left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\boldsymbol{D}\boldsymbol{X}\boldsymbol{V}^{-1/2}\right\|_2 \tag{95}$$

$$=\left\|\sum_{i=1}^n \varepsilon_i v_iv_i^\top\right\|_2=\left\|\sum_{i=1}^n\frac{\varepsilon_i}{\sqrt{f(x_i^\top\phi_*)}}\cdot\Big(\sqrt{f(x_i^\top\phi_*)}\boldsymbol{A}_i\Big)\right\|_2. \tag{96}$$

Note that $\frac{\varepsilon_i}{\sqrt{f(x_i^\top\phi_*)}}$ is a standard normal random variable, for all i. We also define a $d\times d$ matrix $\boldsymbol{\Sigma}=\sum_{i=1}^n f(x_i^\top\phi_*)\boldsymbol{A}_i^2$. Then, we have

$$\boldsymbol{\Sigma}=\sum_{i=1}^n f(x_i^\top\phi_*)\Big(v_iv_i^\top\Big)\Big(v_iv_i^\top\Big) \tag{97}$$

$$=\sum_{i=1}^n f(x_i^\top\phi_*)\left\|v_i\right\|_2^2 v_iv_i^\top. \tag{98}$$

We also know

$$\left\|\sum_{i=1}^n\boldsymbol{A}_i\right\|_2=\left\|\sum_{i=1}^n v_iv_i^\top\right\|_2 \tag{99}$$

$$=\left\|\Big(\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\Big)\Big(\boldsymbol{X}\boldsymbol{V}^{-1/2}\Big)\right\|_2 \tag{100}$$

$$\le\left\|\Big(\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\Big)\right\|_2\left\|\Big(\boldsymbol{X}\boldsymbol{V}^{-1/2}\Big)\right\|_2\le 1, \tag{101}$$

where (101) follows from Lemma A.2. Moreover, we know

$$\left\|\boldsymbol{\Sigma}\right\|_2=\left\|\sum_{i=1}^n f(x_i^\top\phi_*)\left\|v_i\right\|_2^2 v_iv_i^\top\right\|_2 \tag{102}$$

$$\le\left\|d\cdot\sigma^2_{\max}\sum_{i=1}^n v_iv_i^T\right\|_2 \tag{103}$$

$$=d\cdot\sigma^2_{\max}\left\|\sum_{i=1}^n\boldsymbol{A}_i\right\|\le d\cdot\sigma^2_{\max}, \tag{104}$$

where (103) follows from Lemma A.2-A.3, $f(x_i^\top\phi_*)\le\sigma^2_{\max}$, and that $v_iv_i^\top$ is positive semi-definite, and the last inequality follows directly from (101). By Lemma A.4 and the fact that $\varepsilon(x_1),\cdots,\varepsilon(x_n)$ are mutually independent given the contexts $\{x_i\}_{i=1}^n$, we know that

$$\mathbb{P}\Big\{\lambda_{\max}\Big(\sum_{i=1}^n\varepsilon_i\boldsymbol{A}_i\Big)\ge\sqrt{2\left\|\boldsymbol{\Sigma}\right\|_2 s}\Big\}\le d\cdot e^{-s}. \tag{105}$$

Therefore, by choosing $s=\ln(d/\delta)$ and the fact that $\lambda_{\max}\Big(\sum_{i=1}^n\varepsilon_i\boldsymbol{A}_i\Big)=\left\|\sum_{i=1}^n\varepsilon_i\boldsymbol{A}_i\right\|_2$, we obtain

$$\mathbb{P}\Big\{\left\|\sum_{i=1}^n\varepsilon_i\boldsymbol{A}_i\right\|_2\ge\sqrt{2\sigma^2_{\max}d\ln(\frac{d}{\delta})}\Big\}\le\delta. \tag{106}$$

Finally, by applying Lemma 1 and (106) to (94), we conclude that for any $n\in\mathbb{N}$, for any $\delta>0$, with probability at least $1-\delta$, we have

$$\left\|\boldsymbol{X}_n^\top\big(\varepsilon\circ\boldsymbol{X}_n(\theta_*-\widehat{\theta}_n)\big)\right\|_{\boldsymbol{V}_n^{-1}}\le\alpha_n^{(1)}(\delta)\cdot\alpha^{(3)}(\delta). \tag{107}$$

□

A.4. Proof of Lemma 5

We first introduce a useful lemma on the norm of the Hadamard product of two matrices.

Lemma A.5 *Given any two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ of the same dimension, the following holds:*

$$\|\boldsymbol{A} \circ \boldsymbol{B}\|_F \leq \operatorname{tr}(\boldsymbol{A}\boldsymbol{B}^\top) \leq \|\boldsymbol{A}\|_2 \cdot \|\boldsymbol{B}\|_2 \,, \tag{108}$$

where $\|\cdot\|$ denotes the Frobenius norm. When $\boldsymbol{A}$ and $\boldsymbol{B}$ are vectors, the above degenerates to

$$\|\boldsymbol{A} \circ \boldsymbol{B}\|_2 \leq \|\boldsymbol{A}\|_2 \cdot \|\boldsymbol{B}\|_2 \,. \tag{109}$$

Proof of Lemma 5. To simplify notation, we use $\boldsymbol{X}$ and $\boldsymbol{V}$ as a shorthand for $\boldsymbol{X}_n$ and $\boldsymbol{V}_n$, respectively. Let $\boldsymbol{M}$ be a positive definite matrix. We have

$$\|\boldsymbol{A}v\|_{\boldsymbol{M}} = \left\|\boldsymbol{M}^{1/2}\boldsymbol{A}v\right\|_2 \leq \left\|\boldsymbol{M}^{1/2}\boldsymbol{A}\right\|_2 \cdot \|v\|_2 \,, \tag{110}$$

where the last inequality holds since ℓ_2-norm is sub-multiplicative. Meanwhile, we also observe that

$$\left(\theta_* - \widehat{\theta}\right)^\top \boldsymbol{X}^\top \boldsymbol{X}\left(\theta_* - \widehat{\theta}\right) \tag{111}$$

$$= \left(\theta_* - \widehat{\theta}\right)^\top \boldsymbol{V}^{1/2}\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top \boldsymbol{X}\boldsymbol{V}^{-1/2}\boldsymbol{V}^{1/2}\left(\theta_* - \widehat{\theta}\right) \tag{112}$$

$$= \left\|\left(\theta_* - \widehat{\theta}\right)^\top \boldsymbol{V}^{1/2}\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\right\|_2^2 \tag{113}$$

$$\leq \left\|\left(\theta_* - \widehat{\theta}\right)^\top \boldsymbol{V}^{1/2}\right\|_2^2 \left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\right\|_2^2 \tag{114}$$

$$\leq \left\|\theta_* - \widehat{\theta}\right\|_{\boldsymbol{V}}^2 . \tag{115}$$

Therefore, we know

$$\left\|\boldsymbol{X}^\top\left(\boldsymbol{X}\left(\theta_* - \widehat{\theta}\right) \circ \boldsymbol{X}\left(\theta_* - \widehat{\theta}\right)\right)\right\|_{\boldsymbol{V}^{-1}} \tag{116}$$

$$\leq \left\|\boldsymbol{V}^{-1/2}\boldsymbol{X}^\top\right\|_2 \left\|\left(\boldsymbol{X}\left(\theta_* - \widehat{\theta}\right) \circ \boldsymbol{X}\left(\theta_* - \widehat{\theta}\right)\right)\right\|_2 \tag{117}$$

$$\leq 1 \cdot \left\|\boldsymbol{X}\left(\theta_* - \widehat{\theta}\right)\right\|_2^2 \tag{118}$$

$$\leq 1 \cdot \left(\left(\theta_* - \widehat{\theta}\right)^\top \boldsymbol{X}^\top \boldsymbol{X}\left(\theta_* - \widehat{\theta}\right)\right) \tag{119}$$

$$\leq \left\|\theta_* - \widehat{\theta}\right\|_{\boldsymbol{V}}^2 \leq (\alpha_n^{(1)}(\delta))^2, \tag{120}$$

where (118) follows from Lemma A.2 and A.5, and (120) follows from Lemma 1. The proof is complete. ☐

A.5. Proof of Theorem 2

Recall that $h_\beta(u,v) = \left(\Phi\left(\frac{\beta-u}{\sqrt{f(v)}}\right)\right)^{-1}$. We first need the following lemma about Lipschitz smoothness of the function $h_\beta(u,v)$.

Lemma A.6 *The function $h_\beta(u,v)$ defined in (31) is (uniformly) Lipschitz smooth on its domain, i.e., there exists a finite $M_h > 0$ (M_h is independent of u, v, and β) such that for any β with $|\beta| \leq B$, for any $u_1, u_2 \in [-1,1]$ and $v_1, v_2 \in [\sigma^2_{\min}, \sigma^2_{\max}]$,*

$$|\nabla h_\beta(u_1,v_1) - \nabla h_\beta(u_2,v_2)| \leq M_h \left\|\begin{pmatrix}u_1\\v_1\end{pmatrix} - \begin{pmatrix}u_2\\v_2\end{pmatrix}\right\|_2 . \tag{121}$$

Moreover, we have

$$h_\beta(u_2,v_2) - h_\beta(u_1,v_1) \leq \tag{122}$$

$$\begin{pmatrix}u_2-u_1\\v_2-v_1\end{pmatrix}^\top \nabla h_\beta(u_1,v_1) + \frac{M_h}{2}\left\|\begin{pmatrix}u_2-u_1\\v_2-v_1\end{pmatrix}\right\|_2^2 . \tag{123}$$

Proof of Lemma A.6. First, it is easy to verify that $h_\beta(\cdot,\cdot)$ is twice continuously differentiable on its domain $[-1,1] \times [\sigma^2_{\min}, \sigma^2_{\max}]$ and therefore is Lipschitz smooth, for some finite positive constant M_h. To show that there exists an M_h that is independent of u, v, β, we need to consider the gradient and Hessian of $h_\beta(\cdot,\cdot)$. Since $h_\beta(u,v)$ is a composite function that involves $\Phi(\cdot)$ and $f(\cdot)$, it is straightforward to write down the first and second derivatives of $h_\beta(u,v)$ with respect to u and v, which depend on $\Phi(\cdot)$, $\Phi'(\cdot)$, $\Phi''(\cdot)$, $f(\cdot)$, $f'(\cdot)$, and $f''(\cdot)$. Given the facts that for all the u, v and β in the domain of interest, we have $\Phi(\frac{\beta-u}{v}) \in [\Phi(\frac{-B-1}{\sigma^2_{\min}}), 1]$, $\Phi'(\frac{\beta-u}{v}) \in (0, \frac{1}{\sqrt{2\pi}})$, $|\Phi''(\frac{\beta-u}{v})| \leq \frac{B+1}{\sigma_{\min}\sqrt{2\pi}}$, and that $f(\cdot), f'(\cdot), f''(\cdot)$ are all bounded, it is easy to verify that such an M_h indeed exists by substituting the above conditions into the first and second derivatives of $h_\beta(u,v)$ with respect to u and v. Moreover, by Lemma 3.4 in (Bubeck et al., 2015), we know that (123) indeed holds. ☐

Proof of Theorem 2. Define

$$q_u := \sup_{u_0 \in (-1,1)} \left|\frac{\partial h_\beta}{\partial u}\right|\Big|_{u=u_0}, \tag{124}$$

$$q_v := \sup_{v_0 \in (\sigma^2_{\min}, \sigma^2_{\max})} \left|\frac{\partial h_\beta}{\partial v}\right|\Big|_{v=v_0}. \tag{125}$$

By the discussion in the proof of Lemma A.6, we know that q_u and q_v are both positive real numbers. By substituting $u_1 = \theta_1^\top x$, $u_2 = \theta_2^\top x$, $v_1 = f(\phi_1^\top x)$, and $v_2 = f(\phi_2^\top x)$ into (123), we have

$$h_\beta\left(\theta_2^\top x, \phi_2^\top x\right) - h_\beta\left(\theta_1^\top x, \phi_1^\top x\right) \tag{126}$$

$$\leq \begin{pmatrix}(\theta_2-\theta_1)^\top x\\ f(\phi_2^\top x) - f(\phi_1^\top x)\end{pmatrix}^\top \nabla h_\beta(\theta_1^\top x, f(\phi_1^\top x)) \tag{127}$$

$$+ \frac{M_h}{2}\left\|\begin{pmatrix}(\theta_2-\theta_1)^\top x\\ f(\phi_2^\top x) - f(\phi_1^\top x)\end{pmatrix}\right\|_2^2 \tag{128}$$

$$\leq \big(q_u \left\|\theta_2-\theta_1\right\|_{\boldsymbol{M}} \cdot \|x\|_{\boldsymbol{M}^{-1}} \tag{129}$$

$$+ q_v M_f \left\|\phi_2-\phi_1\right\|_{\boldsymbol{M}} \cdot \|x\|_{\boldsymbol{M}^{-1}}\big) \tag{130}$$

$$+ \frac{M_h}{2}\big(\left\|\theta_2-\theta_1\right\|_{\boldsymbol{M}}^2 + M_f^2 \left\|\phi_2-\phi_1\right\|_{\boldsymbol{M}}^2\big) \cdot \|x\|_{\boldsymbol{M}^{-1}} \tag{131}$$

$$\leq (q_u + M_h) \left\|\theta_2-\theta_1\right\|_{\boldsymbol{M}} \cdot \|x\|_{\boldsymbol{M}^{-1}} \tag{132}$$

$$+ M_f(q_v + M_h M_f L) \left\|\phi_2-\phi_1\right\|_{\boldsymbol{M}} \cdot \|x\|_{\boldsymbol{M}^{-1}}, \tag{133}$$

where (130)-(131) follow from the Cauchy-Schwarz inequality and the fact that $f(\cdot)$ is Lipschitz continuous, and (132)-(133) follow from the facts that $\|x\|_2 \leq 1$, $\|\theta_2-\theta_1\|_2 \leq 2$, and $\|\phi_2-\phi_1\|_2 \leq 2L$. By letting $C_3 = q_u + M_h$ and $C_4 = M_f(q_v + M_h M_f L)$, we conclude (32)-(33) indeed holds with C_3 and C_4 being independent of $\theta_1, \theta_2, \phi_1, \phi_2$, and β. ☐

A.6. Proof of Lemma 6

Proof. By Theorem 2 and (35), we know

$$Q_{t+1}^{\text{HR}}(x) - h_{\beta_{t+1}}(\theta_*^\top x, \phi_*^\top x) \tag{134}$$

$$= h_{\beta_{t+1}}(\widehat{\theta}_t^\top x, \widehat{\phi}_t^\top x) + \xi_t(\delta)\|x\|_{\boldsymbol{V}_t^{-1}} - h_{\beta_{t+1}}(\theta_*^\top x, \phi_*^\top x) \tag{135}$$

$$\leq 2\xi_t(\delta)\|x\|_{\boldsymbol{V}_t^{-1}}. \tag{136}$$

Similarly, by switching the roles of $\theta_*^\top, \phi_*^\top$ and $\widehat{\theta}_t^\top, \widehat{\phi}_t^\top$ in (135), we have

$$Q_{t+1}^{\text{HR}}(x) - h_{\beta_{t+1}}(\theta_*^\top x, \phi_*^\top x) \geq 0. \tag{137}$$

☐

A.7. Proof of Theorem 3

Proof. For each user t, let $\pi_t^{\text{HR}} = \{x_{t,1}, x_{t,2}, \cdots\}$ denote the action sequence under the HR-UCB policy. Under HR-UCB, $\widehat{\theta}_t$ and $\widehat{\phi}_t$ are updated only after the departure of each user. This fact implies that $x_{t,i} = x_{t,j}$, for all i, j. Therefore, we can use x_t to denote the action chosen by HR-UCB for the user t, to simplify notation. Let $\overline{R}_t^{\text{HR}}$ denote the expected lifetime of user t under HR-UCB. Similar to (30), we have

$$\overline{R}_t^{\text{HR}} = \Big(\Phi\Big(\frac{\beta_t - \theta_*^\top x_t}{\sqrt{f(\phi_*^\top x_t)}}\Big)\Big)^{-1} = h_{\beta_t}(\theta_*^\top x_t, \phi_*^\top x_t). \tag{138}$$

Recall that π^{oracle} and x_t^* denote the oracle policy and the context of the action of the oracle policy for user t, respectively. We compute the pseudo regret of HR-UCB as

$$\text{Regret}_T = \sum_{t=1}^T \overline{R}_t^* - \overline{R}_t^{\text{HR}} \tag{139}$$

$$= \sum_{t=1}^T h_{\beta_t}\big(\theta_*^\top x_t^*, \phi_*^\top x_t^*\big) - h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big). \tag{140}$$

To simplify notation, we use w_t as a shorthand for $h_{\beta_t}\big(\theta_*^\top x_t^*, \phi_*^\top x_t^*\big) - h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big)$. Given any $\delta > 0$, define an event E_δ in which (12) and (17) hold under the given δ, for all $t \in \mathbb{N}$. By Lemma 1 and Theorem 1, we know that the event E_δ occurs with probability at least $1-3\delta$. Therefore, with probability at least $1-3\delta$, for all $t \in \mathbb{N}$,

$$w_t \leq Q_t^{\text{HR}}(x_t^*) - h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big) \tag{141}$$

$$\leq Q_t^{\text{HR}}(x_t) - h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big) \tag{142}$$

$$= h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big) + \xi_{t-1}(\delta)\|x_t\|_{\boldsymbol{V}_{t-1}^{-1}} \tag{143}$$

$$- h_{\beta_t}\big(\theta_*^\top x_t, \phi_*^\top x_t\big) \tag{144}$$

$$\leq 2\xi_{t-1}(\delta) \cdot \|x_t\|_{\boldsymbol{V}_{t-1}^{-1}}, \tag{145}$$

where (141) and (143) follow directly from the definition of the UCB index, (142) follows from the design of HR-UCB algorithm, and (145) is a direct result under the event E_δ. Now, we are ready to conclude that with probability at least $1-3\delta$, we have

$$\text{Regret}_T = \sum_{t=1}^T w_t \leq \sqrt{T\sum_{t=1}^T w_t^2} \tag{146}$$

$$\leq \sqrt{4\xi_T^2(\delta)T\sum_{t=1}^T \min\{\|x_t\|_{\boldsymbol{V}_{t-1}^{-1}}^2, 1\}} \tag{147}$$

$$\leq \sqrt{8\xi_T^2(\delta)T \cdot d\log\Big(\frac{\mathcal{S}(T)+\lambda d}{\lambda d}\Big)}, \tag{148}$$

where (146) follows from the Cauchy-Schwarz inequality, (147) follows from the fact that $\xi_t(\delta)$ is an increasing function in t, and (148) follows from Lemma 10 and 11 in (Abbasi-Yadkori et al., 2011) and the fact that $\boldsymbol{V}_t = \lambda \boldsymbol{I}_d + \boldsymbol{X}_t^\top \boldsymbol{X}_t = \lambda \boldsymbol{I}_d + \sum_{i=1}^t x_i x_i^\top$. By substituting $\xi_T(\delta)$ into (148) and using the fact that $\mathcal{S}(T) \leq \Gamma(T)$, we know

$$\text{Regret}_T = \mathcal{O}\Bigg(\sqrt{T\log\Gamma(T)\cdot\Big(\log\big(\Gamma(T)\big) + \log(\frac{1}{\delta})\Big)^2}\Bigg). \tag{149}$$

By choosing $\Gamma(T) = KT$ for some constant $K > 0$, we thereby conclude that

$$\text{Regret}_T = \mathcal{O}\Bigg(\sqrt{T\log T\cdot\Big(\log T + \log(\frac{1}{\delta})\Big)^2}\Bigg). \tag{150}$$

The proof is complete. ☐