A Appendix

A.1 Algorithmic Details

We present a precise version of the algorithm described in Section 3 as Algorithm 1. For ease of exposition, we introduce the concept of matrix sampling, which is a notational tool to represent the sampled entries from different subsets of arms in a structured manner.

A.1.1 Matrix Sampling

Consider the $L \times K$ reward matrix $\mathbf{U}$. Consider a 'sampling matrix' $\mathbf{G}$ with dimensions $K \times p$. Let $\{a_1, a_2 \dots a_p\} \subset [K]$. In this work, we consider $\mathbf{G}$ only of the following form: $\mathbf{G}_{a_i,i} = 1,\ \forall 1 \le i \le p$ and zero otherwise. Consider the product between a row s of $\mathbf{U}$ and $\mathbf{G}$, i.e. $\mathbf{U}_{s,:}\mathbf{G}$. This selects the co-ordinates corresponding to $\{a_1 \dots a_p\}$ in vector $\mathbf{U}_{s,:}$. Given a row s (a context s) of $\mathbf{U}$, i.e. $\mathbf{U}_{s,:} := \mathbf{u}[s]$, we describe how to obtain a random Bernoulli vector estimate $\hat{\mathbf{u}}[s]$ such that $\mathbb{E}[\hat{\mathbf{u}}[s]] = \frac{1}{p}\mathbf{U}_{s,:}$ by sampling an arm as follows:

- Given that the context is s, sample a uniform random variable κ with support $\{a_1 \dots a_p\}$, which represents the arm to be pulled after observing the context.
- Conditioned on $\kappa = k$, pull arm k and observe the reward $Y_k \in \{0, 1\}$.
- The random vector sample is then given by $\hat{\mathbf{u}}[s]_k = Y_\kappa \mathbf{e}_\kappa$.

Then we have $\mathbb{E}[\hat{\mathbf{u}}[s]_k] = \mathbb{E}[\mathbb{E}[Y_k|\kappa = k]] = \frac{1}{p}\mathbf{u}[s]_k$. In other words, whenever the context is s, we pull an arm uniformly at random from $\{a_1, a_2 \dots a_p\}$ and the samples are collected in $\hat{\mathbf{u}}[s]$.

A.1.2 Arms to be sampled during *explore*

Before we present the pseudocode, we define the sampling matrices $\{\mathbf{G}(0), \mathbf{G}(1), \cdots, \mathbf{G}(l+1)\}$. Recall that any subset of arms can be encoded in a sampling matrix. $\mathbf{G}(0)$ corresponds to the subset S in Step 1 of *explore* stated in Section 3. For ease of reference, we restate the sets relevant to the context specific sampling procedure in Step 2 of *explore*. $\mathbf{G}(i)$ corresponds to the subset $R(s_t)$ is $s_t \in S(i)$. Let $l = \lfloor K/m \rfloor$ and $r = K\mathrm{mod}(m)$. A set $R \subset [L]$ of contexts is sampled at random, such that $|R| = 2(l+1)m'$ at the onset of the algorithm. We partition R into $l+1$ contiguous subsets $\{S(1), S(2), ..., S(l+1)\}$ of size $2m'$ each. The elements of the set $S(j)$ will be denoted as $S(j) = \{s_1(j), s_2(j) \cdots, s_{2m'}(j)\}$. In Step 2 of *explore*, if $s_t \in S(i)$, then $R(s_t) = \{(i-1)m, (i-1)m + 1, \cdots \max(im-1, K)\}$. If $s_t \notin S(i)$ for all $i \in [l+1]$, then the algorithm is allowed to pull any arm at random, and these samples are ignored.

1. $\mathbf{G}(0)$: An $K \times 2m'$ random matrix formed as follows: An $2m'$ subset $a_1, a_2 \dots a_{2m'} \subset [K]$ is chosen randomly uniformly among all $2m'$-subsets of $[K]$ and $\mathbf{G}(0)_{a_i,i} = 1,\ \forall 1 \le i \le 2m'$ and all other entries are 0.

2. $\mathbf{G}(i)$: An $K \times m$ matrix such that,
$$G(i)_{kj} = \left\{ \begin{array}{ll} 1, & \text{if } k = (i-1)m + j \text{ for } j \in \{1, \cdots, m\} \\ 0, & \text{otherwise} \end{array} \right\}$$
when $i \in \{1, 2, \cdots, l\}$.

3. $\mathbf{G}(l+1)$: An $K \times r$ matrix defined as follows:
$$G(l+1)_{kj} = \left\{ \begin{array}{ll} 1, & \text{if } k = (lm + j) \text{ for } j \in \{1, \cdots r\} \\ 0, & \text{otherwise} \end{array} \right\}$$

In words, $\mathbf{G}(i)$ for $i \in [l]$ is the $K \times m$ matrix which has an identity matrix $I_{m \times m}$ embedded between rows $(i-1)m$ and $im - 1$, and is zero everywhere else.

A.1.3 Representation of the collected Samples

In what follows, let the mean of samples collected through $\mathbf{G}(0)$ till time t be collected in a $L \times 2m'$ matrix $\hat{\mathbf{F}}'(t)$ such that $\mathbb{E}\left[\hat{\mathbf{F}}'(t)\right] = (1/2m')\mathbf{F} = (1/2m')\mathbf{U}\mathbf{G}(0)$ as detailed in Section A.1.1. Let $\hat{\mathbf{F}}(t) = 2m'\hat{\mathbf{F}}'(t)$. Let the samples collected from $\mathbf{G}(i)$ be stored in a $2m' \times m$ matrix $\hat{\mathbf{M}}'_i(t)$ such that $\mathbb{E}\left[\hat{\mathbf{M}}'_i(t)\right] = \frac{1}{m}\mathbf{A}_{S(i),:}\mathbf{W}\mathbf{G}(i)$ for all $i \in \{1, 2, ..., l+1\}$. Let $\hat{\mathbf{M}}_i(t) = m\hat{\mathbf{M}}'_i(t)$ be the scaled version.

A.1.4 Pseudocode

We present a detailed pseudo-code of our algorithm as Algorithm 1. For the sake of completeness we include the robust version of the Hottopix algorithm [18] which is used as a sub-program in Algorithm 1. The following LP is fundamental to the Hottopix algorithm,

$$\begin{aligned} &\min_{\mathbf{C} \in \mathbb{R}_+^{f \times n}} \mathbf{p}^T \mathrm{diag}(\mathbf{C}) \\ &\text{s.t. } \left\| \tilde{\mathbf{X}} - \mathbf{C}\tilde{\mathbf{X}} \right\|_{\infty,1} \le 2\epsilon \\ &\text{and } C_{ii} \le 1,\ C_{ji} \le C_{ii}\ \forall i,j \in [L] \end{aligned} \tag{3}$$

where $\mathbf{p}$ is a vector with distinct positive values.

Algorithm 1 NMF-Bandit - An ϵ-greedy algorithm for Latent Contextual Bandits

At time t,
Observe context $S_t = s_t$
Let $\mathsf{E}(t) \sim \mathrm{Ber}(\epsilon_t)$
if $\mathsf{E}(t) = 1$ **then**
    Explore: Let $H_t \sim \left\{ \begin{array}{cc} \mathrm{Ber}\left(\frac{2m'}{r+2m'}\right), & \text{if } s_t \in S(l+1) \\ \mathrm{Ber}\left(\frac{2m'}{m+2m'}\right), & \text{otherwise} \end{array} \right\}.$
    If $H_t = 1$ sample an arm according to the matrix sampling technique applied to matrix $\mathbf{G}(0)$ and update $\hat{\mathbf{F}}(t)$.
    If $H_t = 0$ sample an arm according to the matrix sampling technique applied to matrix $\mathbf{G}(i)$ if $s_t \in S(i)$ for $i \in \{1, 2, \cdots, l+1\}$ and update $\hat{\mathbf{M}}_i(t)$. If s_t is not in any of these sets then choose an arm at random.
else
    Exploit:
    Let us compute,

$$\hat{\mathbf{W}}(t) = \mathrm{Hottopix}(\mathbf{F}(t), m, 2m'\gamma(t)).$$
$$\hat{\mathbf{A}}(t) = \underset{\mathbf{Z} \geq \mathbf{0}, \mathrm{rowsum}(\mathbf{Z}) = \mathbf{1}}{\mathrm{argmin}} \left\| \mathbf{F}(t) - \mathbf{Z}\hat{\mathbf{W}}(t) \right\|_{\infty,1}.$$

    Let $\hat{\mathbf{W}}(t) \in \mathbb{R}^{m \times K}$ be such that,

$$\hat{\mathbf{W}}(t)_{:,(i-1)m:im-1} = \underset{\mathbf{X}_{m \times m}}{\mathrm{argmin}} \left\| \hat{\mathbf{A}}(t)_{S(i),:}\mathbf{X} - \hat{\mathbf{M}}_i(t) \right\|_2, \ \forall i \in \{1, 2, .., l\}$$
$$\hat{\mathbf{W}}(t)_{:,lm:K} = \underset{\mathbf{X}_{m \times r+1}}{\mathrm{argmin}} \left\| \hat{\mathbf{A}}(t)_{S(l),:}\mathbf{X} - \hat{\mathbf{M}}_{l+1}(t) \right\|_2$$

    Compute $\hat{\mathbf{U}}(t) = \hat{\mathbf{A}}(t)\hat{\mathbf{W}}(t)$. Play the arm a_t such that,

$$a_t = \arg\max_a \hat{\mathbf{U}}(t)_{s_t,a}$$

end if

Algorithm 2 Hottopix($\tilde{\mathbf{X}}, m, \epsilon$)

Input : $\tilde{\mathbf{X}}$ such that $\tilde{\mathbf{X}} = \mathbf{AW} + \mathbf{N}$, where $\mathbf{A} \in [0,1]^{L \times m}$ and $\|A_{i,:}\|_1 = 1$ for all $i \in [L]$, $\mathbf{W} \in \mathbb{R}_+^{m \times 2m'}$ and $\|\mathbf{N}\|_{\infty,1} \leq \epsilon$.
Output : $\hat{\mathbf{W}}$ such that $\hat{\mathbf{W}} \sim \mathbf{W}$.
Compute an optimal solution $\mathbf{C}^*$ to (3).
Let $\mathcal{K}$ denote the set of indices i for which $C^*_{ii} \geq \frac{1}{2}$.
Set $\hat{\mathbf{W}} = \tilde{X}_{\mathcal{K},:}$.

A.2 Theoretical Insights

Below, we discuss some of the key challenges in the theoretical analysis.

Noise Guarantees for samples used in NMF: Matrix completion algorithms that work under the incoherence assumptions require the noise in each element of the matrix to be $O(1/K)$ in order to provide l_∞-norm guarantees on the recovered matrix [20]. In order to ensure such noise guarantees, we require a very large number of samples in order for estimates to concentrate. This in turn increases bandit exploration which implies that regret scales as $O(LK\log(T))$. To avoid this, we follow a different route. In Step 1 of the *explore* phase, the NMF-Bandit algorithm only samples from a small subset of arms denoted by S. By leveraging the ℓ_1-WStRIP property of $\mathbf{W}$, we can ensure that NMF on these samples (which are basically a noisy version of $\mathbf{U}_{:,S}$) gives us a good estimate of $\mathbf{A}$ at time t; this estimate is denoted by $\hat{\mathbf{A}}(t)$. We prove this statement formally in Lemma 6. Given that we sample only from a small subset of arms in the first step of *explore*, in Lemma 11 we show that the samples concentrate sharply enough.

Ensuring enough linear equations to recover W: Recall that the reward matrix has the structure $\mathbf{U} = \mathbf{AW}$. Therefore, an initial approach would be to use the current estimate of $\mathbf{A}$ along with samples of the rewards, and directly recover $\mathbf{W}$. This however will not work due to lack of concentrations. First, the estimate of $\mathbf{A}$ in the early stages will be too noisy to provide sharp estimates about the location of the extreme points aka the latent contexts. Even if we knew the identities of the observed contexts that correspond to "pure" latent contexts (extreme points of the affine space corresponding to the observed contexts), most observed contexts will not correspond to these extreme points – thus, a large number of samples will be wasted, again leading to poor concentrations. Second, if one decides to sample the entries in $\mathbf{U}$ at random, the concentration of the entries would be too weak. As before, these weak concentrations will imply $O(LK\log(T))$ regret.

Instead, we design the context dependent sets of arms to pull in Step 2 of the *explore* phase, such that we get enough independent linear equations to recover $\mathbf{W}$. The key is to have a small number of arms to sample per observed contexts, but the small number of arms differ across observed contexts. In this case, we show that by leveraging the ℓ_2-WStRIP property of $\mathbf{A}$ we can get a good estimate of $\mathbf{W}$, denoted by $\hat{\mathbf{W}}(t)$ even in the presence of sampling noise. Since we sample from a small subset of arms for each observed context, in Lemma 12 we can ensure that we have sharp concentrations.

Scheduling the optimal arm during *exploit*: The l_∞-norm bounds on the errors in $\hat{\mathbf{A}}(t)$ and $\hat{\mathbf{W}}(t)$, imply that $\left\|\hat{\mathbf{U}}(t) - \mathbf{U}\right\|_{\infty,\infty} < \Delta/2$ with probability at least $1 - O(\frac{Lm'}{t})$ provided ϵ_t is sufficiently big (see proof of Theorem 8). Here $\Delta = \min_{s\in[L]}(u^*(s) - \max_{k\neq k^*(s)} U_{s,k})$. This essentially implies that the correct arm is pulled at time t w.h.p if the algorithm decides to *exploit*.

A.3 Description of Generative Models for matrices W and A

The model for $\mathbf{W}$ and $\mathbf{A}$ are both very similar with deterministic and random parts. The technical description of the model given below is complex due to the following two reasons:

1. **Fact 1:** Rows of $\mathbf{A}$ must sum to 1.
2. **Fact 2:** The rows of $\mathbf{W}$ shifted by an arbitrary vector $\mathbf{m} \in \mathbb{R}^{1\times K}$ does not affect the NMF algorithms employed. The setting is invariant to such a shift.

1. *Random+Deterministic Composition*:
   (a) We assume that columns $\mathbf{W}_{:,D}$ corresponding to the column index set $D \subseteq [K]$, $|D| \leq K/(32m)$ is arbitrary and deterministic. $0 \leq W_{i,j} \leq 1,\ j \in D$. The maximum entry in every row of $\mathbf{W}$ is assumed to be contained in the deterministic part.
   (b) Similarly, $\mathbf{A}_{E,:}$ where $E \subseteq [L]$ is arbitrary and deterministic. Let $|E| \leq \rho L$. $\rho = 1/18$. Row sum of every row of $\mathbf{A}_{E,:}$ is 1. In order to ensure separability [33] we assume that there is a subset $M \subseteq E : |M| = m$ such that $\mathbf{A}_{M,:} = \mathbf{I}_{m\times m}$. For all $i \in E - M$, $0 \leq A_{ij} \leq \gamma < 1$.

2. *Bounded randomness in the random part:*

$$\mathbf{W}_{:,D^c} = \mathbf{1} * \mathbf{m}^T + \mathbf{R}_{:,D^c} + \tilde{\mathbf{W}}_{:,D^c} \qquad (4)$$

   (a) (i,j)-th entry of $\tilde{\mathbf{W}}_{:,D^c}$ is an independent mean zero sub-gaussian entry with variance q, and bounded support and sub-gaussian parameter $c(q)$. $\mathbf{m} \in \mathbb{R}^{|D^c|\times 1}$ is an arbitrary deterministic vector [1] .
   (b) $\mathbf{R}_{:,D^c}$ is a deterministic perturbation matrix satisfying $\|\mathbf{R}_{:,j}\|_2 \leq \frac{1}{5},\ \forall j \in D^c$. The support parameters for $\tilde{\mathbf{W}}_{:,D^c}$, $\mathbf{m}$ and $\mathbf{R}_{:,D^c}$ are chosen such that $0 \leq W_{i,j} \leq 1$ a.s. , $\forall j \in D^c$

[1] This is introduced to respect Fact 2 in Section A.3

$\mathbf{A}_{E^c,:}$ is a matrix which is a row-normalized version of another random matrix $\tilde{\mathbf{A}}$. We first describe the random model on the $|E^c| \times m$ matrix $\tilde{\mathbf{A}}$. Like in the case for model of $\mathbf{W}$,

$$\tilde{\mathbf{A}} = \mathbf{N} + \hat{\mathbf{A}} \tag{5}$$

(a) $\hat{\mathbf{A}}$ is a matrix with independent mean zero sub-gaussian entries each with variance q, and bounded support and sub-gaussian parameter $c(q)$.

(b) We denote the matrix of means by $\mathbf{N}$ consisting of the parameters n_{ij}. The ℓ_2 norm of every row of $\mathbf{N}$ is at most $\frac{1}{5}$. The support, sub-gaussian parameter and the matrix of means $\mathbf{N}$ are chosen such that $1/m \leq \tilde{A}_{ij} \leq \gamma < 1$ a.s. The stricter condition (in the lower bound) ensures that after normalization by the row sum, $A_{ij} \leq \gamma < 1,\ i \in E^c$.

A.4 Projection onto a Low Dimensional Space

In this section, we will prove some properties of the matrix $\mathbf{F} = \mathbf{U}\mathbf{G}(0) = \mathbf{A}\mathbf{W}\mathbf{G}(0)$ where $\mathbf{G}(0)$ is a $K \times 2m'$ as defined in Section A.1.1. From the definition in Section 2.1, $\mathbf{A}$ contains a $\mathbf{I}_{m\times m}$ sub-matrix corresponding to the rows in $\mathcal{Z}$. Further, the row sum of every row of $\mathbf{A}$ is 1. This means that the rows of $\mathbf{U}$ consists of points in the convex hull of extreme points, i.e. the rows of $\mathbf{W}$, together with the extreme points themselves.

The extreme points in $\mathbf{W}$ are mapped to extreme points in $\mathbf{W}\mathbf{G}(0)$. We also show that the new set of extreme points $\mathbf{W}\mathbf{G}(0)$ also satisfy what is called the simplical property when $\mathbf{W}$ satisfies the assumptions in Section A.3.

When the entries in $\mathbf{W}$ are random and independent bounded random variables as in Section 2.4, we show that ℓ_1 distance of any non-zero vector $\mathbf{a}$ such that $\mathbf{a}^T\mathbf{1} = 0$ is preserved under the map $\mathbf{a}^T\mathbf{W}\mathbf{G}(0)$with high probability over $\mathbf{W}$ for any fixed $\mathbf{G}(0)$. We need some results relating to sub-gaussianity of the matrix $\mathbf{W}$ which we deal with in the next subsection.

A.5 Sub-gaussianity of a matrix with bounded i.i.d random entries

Definition 7. *[16] A random variable X is sub-gaussian with parameter $c > 0$ if $\mathbb{E}[\exp(tX)] \leq \exp(-c^2t^2),\ \forall t \in \mathbb{R}$.*

Definition 8. *[16] A random vector $\mathbf{Y} \in \mathbb{R}^n$ is isotropic if $\mathbb{E}[(\mathbf{Y}^T\mathbf{x})^2] = \mathbb{E}[\mathbf{x}^T\mathbf{x}],\ \forall \mathbf{x} \in \mathbb{R}^n$. It is sub-gaussian with parameter c if the scalar random variable $\mathbf{Y}^T\mathbf{x}$ is sub-gaussian with parameter c for all $\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1$, i.e. $\mathbb{E}[\exp(t(\mathbf{Y}^T\mathbf{x}))] \leq \exp(-ct^2), \forall t \in \mathbb{R},\ \forall \|\mathbf{x}\|_2 = 1$.*

Lemma 1. *[16],[35] Consider a random variable X such that $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$, $|X| \leq b$ **a.s** for some constant $b > 0$. Then, X is sub-gaussian with parameter $\frac{b^2}{2}$. Consider a random vector $\mathbf{Y} \in \mathbb{R}^n$ where each entry is drawn i.i.d from a mean zero, unit variance and a sub-gaussian distribution with parameter c. Then $\mathbf{Y}$ is a sub-gaussian isotropic vector with the same sub-gaussian parameter c*

Remark: The first part is from Theorem 9.9 in [35] while the second part is from Lemma 9.7 from [16].

Lemma 2. *[16] Let $\mathbf{P}$ and $\mathbf{Q}$ be two matrices of the same dimensions. Let $\sigma_{\min}$ and $\sigma_{\max}$ be the largest and smallest singular values of a matrix respectively. Then,*

$$|\sigma_{\min}(\mathbf{P}) - \sigma_{\min}(\mathbf{Q})| \leq \sigma_{\max}(\mathbf{P} - \mathbf{Q}) \tag{6}$$

Let $\mathbf{P} \in \mathbb{R}^{p\times q}$ where $p \geq q$. Then,

$$\sigma_{\max}(\mathbf{P}^T\mathbf{P} - \mathbf{I}_{n\times n}) \leq \delta \Rightarrow \sigma_{\min}(P) \geq \sqrt(1-\delta) \tag{7}$$

Lemma 3. *[16] Consider an $m\times s$ matrix $\mathbf{P}$ with every row being a random independent sub-gaussian isotropic vector with sub-gaussian parameter c. Let $m > s$, then:*

$$\Pr\left(\sigma_{\max}\left(\frac{1}{m}\mathbf{P}^T\mathbf{P} - \mathbf{I}_{s\times s}\right) \geq \delta\right) \leq 2\exp(-\frac{3\tilde{c}}{4}\delta^2 m + \frac{7s}{2}) \tag{8}$$

Further,

$$\begin{aligned} &\Pr\left(\sigma_s(\mathbf{P}) \leq \sqrt{m}\sqrt{(1-\delta)}\right) && (9)\\ &\leq \Pr\left(\sigma_{\max}\left(\frac{1}{m}\mathbf{P}^T\mathbf{P} - \mathbf{I}_{s\times s}\right) \geq \delta\right) \\ &\leq 2\exp(-\frac{3\tilde{c}}{2}\delta^2 m + \frac{7s}{2}) && (10)\end{aligned}$$

Here, $\tilde{c}$ is a constant that depends only on the sub-gaussian parameter c.

Remark: The first result follows from equation (9.15) in [16] and also from combining Lemma 9.8 and Lemma 9.9 in [16]. The second follows from applying Lemma 2

Definition 9 ([33])**.** *Let us consider a matrix $\mathbf{M}$ which is $p \times q$ where $p \leq q$. Let $\mathbf{m}_i \in \mathbb{R}^{1\times p}$ be the i-th row of the matrix $\mathbf{M}$. The matrix M is α-simplical if $\min\limits_{i\in\{1\cdots p\}} \min\limits_{\mathbf{x}\in\mathrm{conv}(\{\mathbf{m}_1\cdots\mathbf{m}_p\}\setminus\mathbf{m}_i)} \|\mathbf{m}_i - \mathbf{x}\|_1 \geq \alpha$. In other words, every row is at least α far away in ℓ_1 distance from the convex hull of other points.*

A.6 Results regarding sub-matrices of W

The following results hold for $\mathbf{W}\mathbf{G}(0)$) since $\mathbf{W}\mathbf{G}(0) = \mathbf{W}_{:,S}$ when $S = \{a_1 \ldots a_{m'}\}$ is the set of column indices associated with $\mathbf{G}(0)$ as in Section A.1.

Theorem 5. *Let $\mathbf{W}$ follow the random generative model in Section 2.4. Let $S \subseteq D^c$. Let $|S| = m' \geq \frac{512}{21\tilde{c}} m \log(eK)$,*

$$\psi_m(\mathbf{W}_{:,S}) \geq \left(\frac{11}{20}\right)\sqrt{m'} \tag{11}$$

with probability at least $1 - \frac{2}{K^{7m/2}}$ over the randomness in $\mathbf{W}$. Here, $\tilde{c}$ is a constant that depends on the sub-gaussian parameter $c(q)$ of the distributions in the generative model in Section 2.4.

Proof. According to the random generative model for $\mathbf{W}$ in Section 2.4, $\mathbf{W}_S = \tilde{\mathbf{W}}_{:,S} + \mathbf{1}\mathbf{m}_S^T + \mathbf{R}_S$. Here, $\tilde{\mathbf{W}}_{:,S}$ has sub-gaussian entries with parameter $c(q)$, since by Lemma 1, all bounded random variables on support $[-1,1]$ with zero mean are sub-gaussian and their sub-gaussian parameter depends on the variance. Let $\mathbf{m}_S$ refer to the vector restricted to co-ordinate in S. Applying Lemma 3 to the sub-gaussian matrix $(m' \times m)$ $\tilde{\mathbf{W}}_{:,\mathbf{S}}$ with $m' \geq \frac{512}{21\tilde{c}} m \log(eK)$ and setting $\delta = 7/16$, we have:

$$\begin{aligned}\Pr\left(\sigma_m(\tilde{\mathbf{W}}_{:,\mathbf{S}}^T) \leq \frac{3}{4}\sqrt{m'}\right) &\leq 2\exp(-\frac{7}{2} m \log(K)) \\ &\leq 2K^{-7m/2}.\end{aligned}$$

Now, applying Lemma 2, we have:

$$\begin{aligned}|\sigma_m(\mathbf{R}_{:,S} + \tilde{\mathbf{W}}_{:,S}) - \sigma_m(\tilde{\mathbf{W}}_{:,S})| &\leq \sigma_{max}(\mathbf{R}_{:,S}) \\ &\leq \|\mathbf{R}_{:,S}\|_F \\ &\leq \frac{1}{5}\sqrt{m'}\end{aligned}$$

Combining the above two equations, we have:

$$\begin{aligned}&\Pr\left(\sigma_m(\tilde{\mathbf{W}}_{:,S} + \mathbf{R}_{:,S}) \leq \left(\frac{3}{4} - \frac{1}{5}\right)\sqrt{m'}\right) \\ &\leq 2\exp(-\frac{7}{2} m \log(K)) \leq 2K^{-7m/2}.\end{aligned}$$

For any fixed set of size $S = m'$, We have the following chain:

$$\begin{aligned}&\inf_{\mathbf{a}\neq 0:\mathbf{a}^T\mathbf{1}=0} \frac{\|\mathbf{a}^T\mathbf{W}_{:,S}\|_2}{\|\mathbf{a}\|_2} \qquad (12)\\ &= \inf_{\mathbf{a}\neq 0:\mathbf{a}^T\mathbf{1}=0} \frac{\|\mathbf{a}^T(\mathbf{1}\mathbf{m}_S^T + \tilde{\mathbf{W}}_{:,S} + \mathbf{R}_{:,S})\|_2}{\|\mathbf{a}\|_2} \\ &\underset{(\mathbf{a}^T\mathbf{1}=0)}{=} \inf_{\mathbf{a}\neq 0::\mathbf{a}^T\mathbf{1}=0} \frac{\|\mathbf{a}^T(\tilde{\mathbf{W}}_{:,S} + \mathbf{R}_{:,S})\|_2}{\|\mathbf{a}\|_2} \\ &\geq \sigma_m(\mathbf{R}_{:,S} + \tilde{\mathbf{W}}_{:,S}) \qquad (13)\end{aligned}$$

□

Theorem 6. *Consider a matrix $\mathbf{W}$ with the generative model in Section 2.4. Let $m' \geq \frac{512}{21\tilde{c}} m \log(eK)$. For any fixed set S of size $2m'$ such that $S_1 = S \bigcap D$, $|S_1| \leq \frac{2m'}{16m}$ we have:*

$$\psi_m^1(\mathbf{W}_{:,S}) = \inf_{\mathbf{a}\neq 0:\mathbf{a}^T\mathbf{1}=0} \frac{\|\mathbf{a}^T\mathbf{W}_{:,S}\|_1}{\|\mathbf{a}\|_1} \geq \left(\frac{13}{60}\right)\frac{\sqrt{15m'}}{\sqrt{8m}} \tag{14}$$

with probability at least $1 - 2K^{-7m/2}$ over the randomness in $\mathbf{W}$. Further, rows of $\mathbf{W}_{:,S}$ is $\psi_m^1(\mathbf{W}_{:,S})$-simplical

Proof. Let $S_2 = S \bigcap D^c$. Here, $|S_2| \geq 2m'(1 - \frac{1}{16m}) \geq \frac{15m'}{8} \geq \frac{512}{21\tilde{c}} m \log(eK)$. The first result follows from the following chain:

$$\begin{aligned}\|\mathbf{a}^T(\mathbf{W}_{:,S})\|_1 &\geq \|\mathbf{a}^T[\mathbf{W}_{:,S_1}\mathbf{W}_{:,S_2}]\|_2 \qquad (15)\\ &\underset{(a)}{\geq} \|\mathbf{a}^T\mathbf{W}_{:,S_2}\|_2 - \|\mathbf{a}^T\mathbf{W}_{:,S_1}\|_2 \\ &\underset{(b)}{\geq} \|\mathbf{a}\|_2\psi_m(\mathbf{W}_{S_2}) - \|\mathbf{a}\|_2\sqrt{m\frac{2m'}{16m}} \qquad (16)\\ &\underset{(c)}{\geq} \|\mathbf{a}\|_1\frac{\sqrt{15m'}}{\sqrt{8m}}\left(\frac{3}{4} - \frac{1}{5} - \frac{1}{\sqrt{15}}\right) \\ &\geq \left(\frac{3}{4} - \frac{8}{15}\right)\frac{\sqrt{15m'}}{\sqrt{8m}} \text{ w.p. } 1 - \frac{2}{K^{7m/2}} \qquad (17)\end{aligned}$$

Justifications of the above chain are: (a)- Triangle inequality for the norm $\|\cdot\|_2$. (b)- Definition of $\psi_m(\cdot)$ and $\|\mathbf{a}^T\mathbf{W}_{S_1}\|_2 \leq \|\mathbf{a}^T\|_2|\mathbf{W}_{S_1}\|_F \leq \sqrt{m|S_1|}\|\mathbf{a}^T\|_2$. (c)- $\|\cdot\|_2 \geq \frac{\|\cdot\|_1}{\sqrt{m}}$ and applying Theorem 5 because $S_2 \subseteq D^c$ and $|S_2| \geq \frac{512}{21\tilde{c}} m \log(eK)$.

For the second part, let us denote $\mathbf{r}^{-i} \in \mathbb{R}^{1\times m}$ to be a vector satisfying $\sum\limits_{k\neq i} r_k^{-i} = -1$, $r_k^{-i} \leq 0\ \forall k \neq i$ and $r_i^{-i} = 1$. It is easy to easy that:

$$\|\mathbf{r}^{-i}\|_1 \geq 1. \tag{18}$$

From the definition for an α-simplical matrix (Definition 9), it is enough to show that for any $\mathbf{r}^{-i}$, $\|\mathbf{r}^{-i}\mathbf{W}_S\|_1 \geq \psi_m^1(\mathbf{W}_{:,S})$. We prove this as follows:

$$\|\mathbf{r}^{-i}\mathbf{W}_{:,S}\|_1 \underset{(\|\mathbf{r}^{-i}\|_1\geq 1)}{\geq} \psi_m^1(\mathbf{W}_{:,S}) \tag{19}$$

□

A.6.1 Choosing a good S for $\mathbf{G}(0)$

Lemma 4. *Let D be the set as defined in Section 2.4. Let a random $2m'$-subset S be chosen out of $[K]$ where $m' = \frac{512}{21\tilde{c}} m \log(eK)$. Then, $\Pr\left(|S \bigcap D| \leq \frac{2m'}{16m}\right) \leq \exp(-c_1 \log(eK))$ for constant $c_1 > 0$ that depends on $\tilde{c}$.*

Proof. Let $X_1, \dots X_{2m'}$ be set of indicator functions such that $X_i = 1$ if the i-th element in the random subset S chosen uniformly without replacement belongs to D and it is 0 otherwise. Let $Y_1, Y_2 \dots Y_{2m'}$ be the set of indicator functions such that $Y_i = 1$ (and 0 otherwise) if the i-th element in the random multiset S belongs to D where the multiset elements are chosen independently and uniformly with replacement. It is clear that $\mathbb{E}[X_i] = \mathbb{E}[Y_i] = \frac{|D|}{K} = \mu \geq \frac{1}{32m}$. The moment generating function of the sum of X_i's is dominated by the moment generating function of the sum of Y_i's. Therefore, all concentration inequalities, based on moment generating functions, for variables drawn with replacement holds for variables drawn without replacement [23]. In particular, the following inequality derived from moment generating functions holds [24] for any $\delta > 0$:

$$\begin{aligned}
&\Pr\left(\sum X_i \geq (1+\delta)2m'\mu\right) \\
&\leq \Pr\left(\sum X_i \geq (1+\delta)2m'\mu\right) \\
&\leq \exp\left(\delta 2m'\mu\right)(1+\delta)^{-(1+\delta)2m'\mu}.
\end{aligned}$$

Let us take $\delta = 1$. Therefore, $\Pr\left(|S \bigcap D| \geq \frac{2m'}{16m}\right) \leq \left(\frac{4}{e}\right)^{-\frac{32}{21\tilde{c}}\log(eK)} \leq \frac{1}{(eK)^{\frac{32}{21\tilde{c}}\log(4/e)}}$.

$\square$

Proof of Theorem 2. From Theorem 6 and Lemma 4 we have,

$$\begin{aligned}
&\mathbb{E}_{\mathbf{W}}\left[\mathbb{P}_{\mathbf{S}}\left(\psi_m^1(\mathbf{W}_{:,S}) < \left(\frac{13}{60}\right)\frac{\sqrt{15m'}}{\sqrt{8m}}\right)\right] \\
&\leq \exp(-c_1\log(eK)) + 2K^{-7m/2} \\
&\leq 2\exp(-c_1\log(eK)
\end{aligned}$$

Now by Markov's inequality this implies that,

$$\begin{aligned}
&\mathbb{P}_{\mathbf{W}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\psi_m^1(\mathbf{W}_{:,S}) < \left(\frac{13}{60}\right)\frac{\sqrt{15m'}}{\sqrt{8m}}\right)\right.\right. \\
&\left.\left.\geq 2\exp(-\frac{c_1}{2}\log(eK))\right)\right] \\
&\leq \frac{\exp(-c_1\log(eK)}{\exp(-\frac{c_1}{2}\log(eK))} \\
&\leq \exp(-\frac{c_1}{2}\log(eK))
\end{aligned}$$

This implies the following chain:

$$\begin{aligned}
&\mathbb{P}_{\mathbf{W}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\psi_m^1(\mathbf{W}_{:,S}) > \left(\frac{13}{60}\right)\frac{\sqrt{15m'}}{\sqrt{8m}}\right)\right.\right. \\
&\left.\left.\leq 1 - 2\exp(-\frac{c_1}{2}\log(eK))\right)\right] \\
&\leq \exp(-\frac{c_1}{2}\log(eK)) \\
&\Rightarrow \mathbb{P}_{\mathbf{W}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\psi_m^1(\mathbf{W}_{:,S}) > \left(\frac{13}{60}\right)\frac{\sqrt{15m'}}{\sqrt{8m}}\right)\right.\right. \\
&\left.\left.\geq 1 - 2\exp(-\frac{c_1}{2}\log(eK))\right)\right] \\
&\geq 1 - \exp(-\frac{c_1}{2}\log(eK))
\end{aligned}$$

This proves that with probability at least $1 - \exp(-\frac{c_1}{2}\log(eK))$ the ℓ_1-WStRIP condition is satisfied with the said parameters. $\square$

A.7 Results regarding sub-matrices of A

We assume that $\mathbf{A}$ satisfies the random generative model in 2.4. We prove some results regarding the minimum singular values of sub-matrices corresponding to columns in set S ($|S| = 2m'$) which is a mix of random and the deterministic columns. The proofs follow closely those of $\mathbf{W}$ in the previous section.

Theorem 7. *Let $\mathbf{A}$ follow the random generative model in Section 2.4. Let $m' \geq \frac{512}{21\tilde{c}}m\log(eL)$. Fix any set S of size $2m'$ such that $S_1 = S \bigcap E$, $|S_1| \leq \frac{2m'}{9}$. Let $S_2 = S \setminus S_1$. Then, we have:*

$$\sigma_m\left(\mathbf{A}_{S,:}\right) \geq \frac{\sqrt{m'}}{m}\left(\frac{1}{20}\right) \text{ w.p } 1 - \frac{2}{L^{7m/2}}. \tag{20}$$

Proof. Let $\tilde{S}_2$ be the set of rows in the random matrix $\tilde{\mathbf{A}}$ that corresponds to the rows S_2 in $\mathbf{A}$. Here, $\hat{\mathbf{A}}_{\tilde{S}_2,:}$ has sub-gaussian entries with sub-gaussian parameter $c(q)$, since by Lemma 1, all bounded random variables on support $[-1, 1]$ with zero mean are sub-gaussian and their sub-gaussian parameter depends on the variance.

Therefore, applying Lemma 3 to the sub-gaussian matrix ($|\tilde{S}_2| \times m$) $\hat{\mathbf{A}}_{\tilde{\mathbf{S}}_2,:}$ with $|\tilde{S}_2| \geq m' \geq \frac{512}{21\tilde{c}}m\log(eL)$ and setting $\delta = 7/16$, we have:

$$\begin{aligned}
\Pr\left(\sigma_m(\hat{\mathbf{A}}_{\tilde{\mathbf{S}}_2,:}) \leq \frac{3}{4}\sqrt{m'}\right) &\leq 2\exp\left(-\frac{7m\log(L)}{2}\right) \\
&\leq 2L^{-7m/2}.
\end{aligned}$$

Now, consider the following matrix: $[\frac{1}{m}\left(\mathbf{N}_{\tilde{S}_2,:} + \hat{\mathbf{A}}_{\tilde{S}_2,:}\right)\mathbf{A}_{S_1,:}]$. First, note that according to the model in Section 2.4, rows of $\mathbf{A}_{S_1,:}$ sum to 1. Therefore, we have the following chain for any non zero vector $\mathbf{a} \in \mathbb{R}^{1\times m}$:

$$\|[\mathbf{N}_{\tilde{S}_2,:} + \hat{\mathbf{A}}_{\tilde{S}_2,:}\mathbf{A}_{S_1,:}]\mathbf{a}\|_2 \geq \|\mathbf{N}_{\tilde{S}_2,:} + \frac{1}{m}\hat{\mathbf{A}}_{\tilde{S}_2,:}\mathbf{a}\|_2 \tag{21}$$

$$\begin{aligned}
&- \|\mathbf{A}_{S_1,:}\mathbf{a}\|_2 \\
&\geq \|\left(\mathbf{N}_{\tilde{S}_2,:} + \hat{\mathbf{A}}_{\tilde{S}_2,:}\right)\mathbf{a}\|_2 - \sqrt{\sum_{i\in S_1}\|\mathbf{A}_{i,:}\|_2^2\|\mathbf{a}\|_2^2} \\
&\geq \|\left(\mathbf{N}_{\tilde{S}_2,:} + \hat{\mathbf{A}}_{\tilde{S}_2,:}\right)\mathbf{a}\|_2 - \sqrt{\sum_{i\in S_1}\|\mathbf{A}_{i,:}\|_1^2\|\mathbf{a}\|_2^2} \\
&\geq \|\left(\mathbf{N}_{\tilde{S}_2,:} + \hat{\mathbf{A}}_{\tilde{S}_2,:}\right)\mathbf{a}\|_2 - \sqrt{2\rho m'}\|\mathbf{a}\|_2 \\
&\geq \|\hat{\mathbf{A}}_{\tilde{S}_2,:}\mathbf{a}\|_2 - \|\mathbf{N}_{\tilde{S}_2,:}\mathbf{a}\|_2 - \sqrt{2\rho m'}\|\mathbf{a}\|_2 \\
&\geq \sigma_m\left(\hat{\mathbf{A}}_{\tilde{S}_2,:} - \sqrt{2m'(1-\frac{1}{9})}\frac{1}{5} - \sqrt{2\frac{1}{9}m'}\right)\|\mathbf{a}\|_2 \\
&\geq \left(\frac{3}{4} - \frac{1}{5} - \frac{\sqrt{2\frac{1}{9}}}{\sqrt{1-\frac{1}{9}}}\right)\sqrt{2m'(1-\frac{1}{9})} \text{ w.p. } 1-2L^{-7m/2}. \\
&\geq \left(\frac{3}{4} - \frac{1}{5} - \frac{1}{2}\right)\sqrt{m'} \text{ w.p. } 1-2L^{-7m/2}.
\end{aligned} \tag{22}$$

Now, we normalize the every row of $[\mathbf{N}_{\tilde{S}_2,:}+\hat{\mathbf{A}}_{\tilde{S}_2,:}\mathbf{A}_{S_1,:}]$ to get $[\mathbf{A}_{S_2,:}\mathbf{A}_{S_1,:}] = \mathbf{A}_S\mathbf{P}$ where $\mathbf{P}$ is a permutation matrix. Now, every entry gets scaled by at least $1/m$ since rows sum is at most m. Therefore, the minimum singular value scales by at least $1/m$. Therefore,

$$\sigma_m(\mathbf{A}_S) = \sigma_m(\mathbf{A}_S\mathbf{P}) \geq \frac{\sqrt{m'}}{m}\left(\frac{3}{4} - \frac{1}{5} - \frac{1}{2}\right) \\ \text{w.p } 1-2L^{-7m/2}.$$

□

A.7.1 Choosing a good $S(i)$ for a $\mathbf{G}(i)$

Lemma 5. *Let E be the set as defined in Section 2.4. Let a random $2m'$-subset S be chosen out of $[L]$ where $m' = \frac{512}{21\tilde{c}}m\log(eL)$. Then, $\Pr\left(|S\bigcap E| \leq \frac{2m'}{9}\right) \leq \exp(-c_2 m\log(eL))$ for constant $c_2 > 0$ that depends on $\tilde{c}$.*

Proof. The proof is identical to the proof of Lemma 4. We just choose $\mu = \frac{1}{18}$ and $\delta = 1$. Therefore we have:

$$\Pr\left(|S\bigcap E| \geq \frac{2m'}{9}\right) \leq \frac{1}{(eL)^{\frac{512\log(4/e)}{189\tilde{c}}m}}. \tag{23}$$

□

Proof of Theorem 3. From Theorem 7 and Lemma 5 we have,

$$\mathbb{E}_{\mathbf{A}}\left[\mathbb{P}_{\mathbf{S}}\left(\sigma_m(\mathbf{A}_{S,:}) < \frac{\sqrt{m'}}{m}\left(\frac{1}{20}\right)\right)\right] \\ \leq 3\exp(-c_2' m\log(eL))$$

Now by Markov's inequality this implies that,

$$\begin{aligned}
&\mathbb{P}_{\mathbf{A}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\sigma_m(\mathbf{A}_{S,:}) < \frac{\sqrt{m'}}{m}\left(\frac{1}{20}\right)\right.\right.\right. \\
&\left.\left.\left.\geq \exp(-\frac{c_2'}{2}m\log(eL))\right)\right)\right] \\
&\leq 3\frac{\exp(-c_2' m\log(eL))}{\exp(-\frac{c_2'}{2}m\log(eL))} \\
&\leq 3\exp(-\frac{c_2'}{2}m\log(eL))
\end{aligned}$$

This implies the following chain:

$$\begin{aligned}
&\mathbb{P}_{\mathbf{A}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\sigma_m(\mathbf{A}_{S,:}) > \frac{\sqrt{m'}}{m}\left(\frac{1}{20}\right)\right.\right.\right. \\
&\left.\left.\left.\leq 1-\exp(-\frac{c_2'}{2}m\log(eL))\right)\right)\right] \\
&\leq 3\exp(-\frac{c_2'}{2}m\log(eL)) \\
&\Rightarrow \mathbb{P}_{\mathbf{A}}\left[\left(\mathbb{P}_{\mathbf{S}}\left(\sigma_m(\mathbf{A}_{S,:}) > \frac{\sqrt{m'}}{m}\left(\frac{1}{20}\right)\right.\right.\right. \\
&\left.\left.\left.\geq 1-\exp(-\frac{c_2'}{2}m\log(eL))\right)\right)\right] \\
&\geq 1-3\exp(-\frac{c_2'}{2}m\log(eL))
\end{aligned}$$

This proves that with probability at least $1-\exp(-\frac{c_2'}{2}m\log(eL))$ the ℓ_2-WStRIP condition is satisfied with the said parameters. □

A.8 Noisy NMF in Low dimensions

In this section we enhance the guarantees of the robust Hottopix algorithm from [18] provided $\mathbf{W}$ satisfies ℓ_1-WStRIP and the subset S chosen by Algorithm 1 is good as in Section 4.

Lemma 6. *Suppose $\mathbf{W}$ satisfies ℓ_1-WStRIP with parameter $(\delta,\rho_1,2m')$ and the subset S of its columns ($|S| = 2m'$) satisfies $\psi_m^1(\mathbf{W}_{S,:}) \geq \rho_1$. Consider a matrix $\tilde{\mathbf{X}} = \mathbf{A}\mathbf{W}_{:,S} + \mathbf{N}$ such that $\|\mathbf{N}\|_{\infty,1} \leq \epsilon$ and $\mathbf{A}$ is separable [33]. Under these assumptions* Hottopix$(\tilde{\mathbf{X}}, m, \epsilon)$ *returns $\hat{\mathbf{W}}$ such that,*

$$\left\|\hat{\mathbf{W}} - \mathbf{W}_{:,S}\right\|_{\infty,1} \leq \epsilon \tag{24}$$

if $\epsilon < \frac{\rho_1(1-\lambda)}{15}$. *Suppose* $\hat{\mathbf{A}} = \mathrm{argmin}_{\mathbf{Z}\geq\mathbf{0},\mathrm{rowsum}(\mathbf{Z})=\mathbf{1}} \left\|\tilde{\mathbf{X}} - \mathbf{Z}\hat{\mathbf{W}}\right\|_{\infty,1}$. *Then we have,*

$$\left\|\hat{\mathbf{A}} - \mathbf{A}\right\|_{\infty,1} \leq \frac{4\epsilon}{\rho_1 - \epsilon} \tag{25}$$

Proof. Let $\mathbf{W}' = \mathbf{W}_{:,S}$ and $\mathbf{X} = \mathbf{A}\mathbf{W}_{:,S}$. The bound in (6) is immediate from Theorem 2 in [18] as $\mathbf{W}'$ is ρ_1-robust simplical by Theorem 6. We first note that,

$$\begin{aligned}\left\|\tilde{\mathbf{X}} - \mathbf{A}\hat{\mathbf{W}}\right\|_{\infty,1} &\leq \left\|\tilde{\mathbf{X}} - \mathbf{X}\right\|_{\infty,1} + \|X - \mathbf{A}\mathbf{W}'\|_{\infty,1} \\ &\quad + \left\|\mathbf{A}\mathbf{W}' - \mathbf{A}\hat{\mathbf{W}}\right\|_{\infty,1} \\ &\leq \left\|\mathbf{A}\left(\mathbf{W}' - \hat{\mathbf{W}}\right)\right\|_{\infty,1} + \epsilon \\ &\leq \|\mathbf{A}\|_{\infty,1}\left\|\mathbf{W}' - \hat{\mathbf{W}}\right\|_{\infty,1} + \epsilon \leq 2\epsilon\end{aligned}$$

The first inequality follows from the triangle inequality while the last one holds because $\|\mathbf{A}\|_{\infty,1} = 1$. Thus, the LP to recover $\hat{\mathbf{A}}$ will always output $\hat{\mathbf{A}}$ with,

$$\left\|\mathbf{X} - \mathbf{A}\hat{\mathbf{W}}\right\|_{\infty,1} = \left\|\mathbf{A}\mathbf{W}' - \hat{\mathbf{A}}\hat{\mathbf{W}}\right\|_{\infty,1} \leq 3\epsilon. \tag{26}$$

We can apply triangle inequality to get,

$$\begin{aligned}\left\|\left(\mathbf{A} - \hat{\mathbf{A}}\right)\mathbf{W}'\right\|_{\infty,1} &\leq \left\|\mathbf{A}\mathbf{W}' - \hat{\mathbf{A}}\hat{\mathbf{W}}\right\|_{\infty,1} \\ &\quad + \left\|\hat{\mathbf{A}}\left(\mathbf{W}' - \hat{\mathbf{W}}\right)\right\|_{\infty,1} \\ &\leq 3\epsilon + \left\|\hat{\mathbf{A}}\right\|_{\infty,1}\left\|\mathbf{W}' - \hat{\mathbf{W}}\right\|_{\infty,1} \\ &\leq 3\epsilon + \left(1 + \left\|\hat{\mathbf{A}} - \mathbf{A}\right\|_{\infty,1}\right)\epsilon\end{aligned} \tag{27}$$

In order to get the desired result we need to lower bound the L.H.S in (27). Note that $\mathrm{rowsum}\left(\mathbf{A} - \hat{\mathbf{A}}\right) = \mathbf{0}$. Therefore we have,

$$\left\|\left(\mathbf{A} - \hat{\mathbf{A}}\right)\mathbf{W}'\right\|_{\infty,1} \geq \left\|\mathbf{A} - \hat{\mathbf{A}}\right\|_{\infty,1}\rho_1 \tag{28}$$

by definition. Combining (28) and (27) we get the required bound. ☐

A.9 Noisy Recovery of Extreme Points

In this section we assume that $\mathbf{A}$ satisfies the ℓ_2-WStRIP property with parameter $(\delta/L, \rho_2, m')$.

Lemma 7. *If* $\mathbf{A}$ *satisfies the* ℓ_2*-WStRIP property with parameter* $(\delta/L, \rho_2, 2m')$ *then the sets* $\{S(1), \cdots, S(l+1)\}$ *with* $|S(i)| = 2m'$ *satisfy,*

$$\sigma_m(\mathbf{A}_{S(i),:}) \geq \rho_2, \;\; \textit{for all } i \in [l+1]$$

with probability atleast $1-\delta$ *over the randomness in choosing the subsets.*

Proof. The proof of this lemma is just an union bound over all the events $\left\{\sigma_m(\mathbf{A}_{S(i),:}) < \rho_2\right\}$. Note that by virtue of ℓ_2-WStRIP each of these events is true with probability atmost δ/L. ☐

If the conditions of the above lemma are satisfied we will call the corresponding sets *good*. Recall the definition of $\hat{\mathbf{M}}_i(t)$. We will show that if $\hat{\mathbf{A}}(t)$ is close to $\mathbf{A}$ and the matrices $\hat{\mathbf{M}}_i(t)$ are sufficiently close to their means, then we recover $\mathbf{W}$ upto the same accuracy. Let us define $\mathbf{M}_i = \mathbb{E}\left[\hat{\mathbf{M}}_i(t)\right]$.

Lemma 8. *Suppose* $\mathbf{A}$ *satisfies the* ℓ_2*-WStRIP property and* $\{S(1), S(2), \cdots S(l+1)\}$ *are good in the sense of Lemma 7. Given that* $\left\|\hat{\mathbf{A}}(t) - \mathbf{A}\right\|_{\infty,1} \leq \epsilon_1$ *and* $\left\|\hat{\mathbf{M}}_i(t) - \mathbf{M}_i\right\|_{\infty,\infty} \leq \epsilon_2$ *for all* $i \in [l+1]$, $\hat{\mathbf{W}}(t)$ *recovered by Algorithm 1 satisfies,*

$$\left\|\hat{\mathbf{W}}(t) - \mathbf{W}\right\|_{\infty,\infty} \leq \frac{m(2\epsilon_1 + 3\epsilon_2)}{\rho_2} \tag{29}$$

if $\epsilon_1, \epsilon_2 \leq \frac{\rho_2}{m}$.

Proof. Let $\hat{\mathbf{W}}(t)_{:,(i-1)m:im-1}$ and $\mathbf{W}_{:,(i-1)m:im-1}$ be denoted by $\hat{\mathbf{W}}_i(t)$ and $\mathbf{W}_i$ respectively. Similarly we denote $\hat{\mathbf{A}}(t)_{S(i),:}$ and $\mathbf{A}_{S(i),:}$ by $\hat{\mathbf{A}}_i(t)$ and $\mathbf{A}_i$ respectively. Then following identities hold,

$$\begin{aligned}\mathbf{A}_i\mathbf{W}_i &= \mathbf{M}_i \\ \hat{\mathbf{A}}_i(t)\hat{\mathbf{W}}_i(t) &= \hat{\mathbf{M}}_i(t)\end{aligned} \tag{30}$$

Note that $\mathbf{A}_i$ has full-column rank. Let the left-inverse of $\mathbf{A}_i$ be $\mathbf{A}_i^*$. It is easy to see that,

$$\|\mathbf{A}_i^*\|_{\infty,1} \leq \frac{m}{\rho_2}. \tag{31}$$

From (30) we have,

$$\begin{aligned}&\left(\mathbf{I} + \mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\right)\hat{\mathbf{W}}_i(t) = \mathbf{W}_i + \mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i) \\ &\implies \hat{\mathbf{W}}_i(t) \\ &= \left(\mathbf{I} + \mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\right)^{-1}\left(\mathbf{W}_i + \mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i)\right) \\ &\implies \hat{\mathbf{W}}_i(t) \\ &= \left(\mathbf{I} - \mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)(\mathbf{I} + \mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i))\right)(\mathbf{W}_i \\ &+ \mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i))\end{aligned}$$

We can simplify further to yield,

$$\begin{aligned}&\hat{\mathbf{W}}_i(t) - \mathbf{W}_i = \mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i) \\ &- \left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\mathbf{W}_i + \left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\right)^2\mathbf{W}_i\right) \\ &- \left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i)\right. \\ &\left.+ \left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t) - \mathbf{A}_i)\right)^2\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t) - \mathbf{M}_i)\right)\end{aligned}$$

Therefore by triangle inequality we have,

$$\left\|\hat{\mathbf{W}}_i(t)-\mathbf{W}_i\right\|_{\infty,1}=\left\|\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right\|_{\infty,1}$$
$$+\left\|\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\mathbf{W}_i+\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\right)^2\mathbf{W}_i\right)\right\|_{\infty,1}$$
$$+\left\|\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right.\right.$$
$$\left.\left.+\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\right)^2\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right)\right\|_{\infty,1}$$

Now we will bound each of the terms seperately as follows,

$$\left\|\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right\|_{\infty,1}\le\|\mathbf{A}_i^*\|_{\infty,1}\left\|(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right\|_{\infty}$$
$$\le\frac{m\epsilon_2}{\rho_2}$$

Similarly we have,

$$\left\|\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\mathbf{W}_i+\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\right)^2\mathbf{W}_i\right)\right\|_{\infty,1}$$
$$\le\|\mathbf{A}_i^*\|_{\infty,1}(1+\|\mathbf{A}_i^*\|_{\infty,1}\epsilon_1)\epsilon_1\|\mathbf{W}_i\|_{\infty,\infty}$$
$$\le\frac{2m\epsilon_1}{\rho_2}$$

Finally the third term can be bounded as,

$$\left\|\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right.\right.$$
$$\left.\left.+\left(\mathbf{A}_i^*(\hat{\mathbf{A}}_i(t)-\mathbf{A}_i)\right)^2\mathbf{A}_i^*(\hat{\mathbf{M}}_i(t)-\mathbf{M}_i)\right)\right\|_{\infty,1}$$
$$\le\left(\|\mathbf{A}_i^*\|_{\infty,1}\right)^2\epsilon_1\epsilon_2+\left(\|\mathbf{A}_i^*\|_{\infty,1}\right)^3\epsilon_1^2\epsilon_2\le\frac{2m\epsilon_2}{\rho_2}$$

Therefore we have,

$$\left\|\hat{\mathbf{W}}_i(t)-\mathbf{W}_i\right\|_{\infty,1}\le\frac{m(2\epsilon_1+3\epsilon_2)}{\rho_2}$$

We can repeat the same analysis for all $i\in[l+1]$ to arrive at the required result. ☐

A.10 Putting it together: Online Analysis

In this section we prove Theorem 8, which provides a parameter dependent upper bound on the regret of Algorithm 1 if $\mathbf{W}$ and $\mathbf{A}$ satisfy the ℓ_1-WStRIP and ℓ_2-WStRIP. The regret bound provided here is in the parameter dependent regime, that is we assume a constant gap between the best arm and the rest for each context. More precisely let $\Delta=\min_{s\in[L]}\left(u^*(s)-\max_{k\neq k*(s)}U_{sk}\right)$ be a fixed constant not scaling with L,K or t. This falls under the purview of the random generative model because we allow for $\Theta(K/m)$ deterministic rewards for each of the latent context. These conditions are expected to hold in real world data as each latent contexts are expected to have some unique arms which are significantly different from the others. In the said regime we reduce the regret bound of $O\left(LK\log(t)\right)$ for general contextual bandit to only an $O\left(L\mathrm{poly}(m,\log(K))\log(T)\right)$ dependence.

Theorem 8. *In a contextual bandit setting suppose the reward matrix has the form $\mathbf{U}=\mathbf{AW}$ and each contexts s arrives independently with probability β_s for all $s\in[L]$. Assume that $L=\Omega(K\log(K))$. If the problem parameters satisfy the following assumptions,*

- $\beta=\min_s\beta_s=\Omega\left(1/L\right)$.
- $\mathbf{W}\in\mathbb{R}^{m\times K}$ *satisfies ℓ_1-WStRIP with parameters $(\delta,\rho_1,2m')$*
- $\mathbf{A}\in[0,1]^{L\times m}$ *satisfies ℓ_2-WStRIP with parameters $(\delta/L,\rho_2,2m')$ and is separable [33].*

then with probability atleast $1-\delta$, Algorithm 1 with $\epsilon_t=\min\left(1,\frac{\theta(2m'+m)}{\beta t}\right)$ and $\gamma(t)=\max\left(\frac{1}{t},\frac{2}{\sqrt{\theta}}\right)$ has regret,

$$R(T)\le\frac{\theta(m+2m')\log(T)}{\beta}+4(L+K+1)m'\log(T)+o(1)$$
$$=O\left(L\frac{\mathrm{poly}(m,m')}{\Delta^2}\log T\right)$$
$$=O\left(L\frac{m^5\log^2K}{\Delta^2}\log T\right)$$

where $\theta\ge4\max\left(\frac{2m'((16+\Delta)\rho_2+32m)}{\Delta\rho_1\rho_2},\frac{15}{\rho_1(1-\lambda)}\right)^2$.

Before we proceed to the proof of our theorem, we need to introduce a few useful lemmas. The next lemma connects the chance of making an error in the *exploit* phase with the estimation errors in the system.

Lemma 9. *Suppose at time t, $\left\|\hat{\mathbf{F}}(t)-\mathbf{F}\right\|_{\infty,\infty}\le\epsilon_1(t)$ and $\left\|\hat{\mathbf{M}}_i(t)-\mathbf{M}_i\right\|_{\infty}\le\epsilon_2(t)$ for all $i\in[l+1]$. If the following conditions hold,*

$$\epsilon_1(t)\le\min\left(\frac{\Delta\rho_1\rho_2}{2m'((16+\Delta)\rho_2+32m)},\frac{\rho_1(1-\lambda)}{15}\right)$$
$$\epsilon_2(t)\le\frac{\Delta\rho_2}{12m}$$
$$E(t)=0\tag{32}$$

then $k(t)=k^*(s_t)$, that is the optimal arm for the context is scheduled in the exploit phase.

Proof. If $\epsilon_1(t)\le\frac{\rho_1(1-\lambda)}{15}$, then by Lemma 6 we have,

$$\left\|\hat{\mathbf{A}}(t)-\mathbf{A}\right\|_{\infty,1}\le\frac{8m'\epsilon_1(t)}{\rho_1-2m'\epsilon_1(t)}\tag{33}$$

Since we have,

$$\epsilon_1(t) \leq \frac{m\rho_1}{2m'(4\rho_2+m)}$$
$$\epsilon_2(t) \leq \frac{\rho_2}{m}$$

it is easy to verify that the conditions of Lemma 8 are satisfied. Therefore we have,

$$\left\|\hat{\mathbf{W}}(t) - \mathbf{W}(t)\right\|_{\infty,\infty} \leq \frac{m}{\rho_2}\left(\frac{16m'\epsilon_1(t)}{\rho_1 - 2m'\epsilon_1(t)} + 3\epsilon_2(t)\right) \tag{34}$$

Therefore we have,

$$\begin{aligned}
&\left\|\hat{\mathbf{U}}(t) - \mathbf{U}\right\|_{\infty,\infty} = \left\|\mathbf{A}\mathbf{W} - \hat{\mathbf{A}}(t)\hat{\mathbf{W}}(t)\right\|_{\infty,\infty} \\
&\leq \|\mathbf{A}\|_{\infty,1}\left\|\mathbf{W} - \hat{\mathbf{W}}(t)\right\|_{\infty,\infty} \\
&+ \left\|\mathbf{A} - \hat{\mathbf{A}}(t)\right\|_{\infty,1}\left\|\hat{\mathbf{W}}(t)\right\|_{\infty,\infty} \\
&\leq \frac{m}{\rho_2}\left(\frac{16m'\epsilon_1(t)}{\rho_1 - 2m'\epsilon_1(t)} + 3\epsilon_2(t)\right) + \frac{8m'\epsilon_1(t)}{\rho_1 - 2m'\epsilon_1(t)} \\
&\leq \frac{8m'\epsilon_1(t)}{\rho_1 - 2m'\epsilon_1(t)}\left(1 + \frac{2m}{\rho_2}\right) + 3\frac{m\epsilon_2(t)}{\rho_2}
\end{aligned}$$

Now, under the conditions of the lemma in (32), we have

$$\frac{8m'\epsilon_1(t)}{\rho_1 - 2m'\epsilon_1(t)}\left(1 + \frac{2m}{\rho_2}\right) \leq \frac{\Delta}{4}$$
$$3\frac{m\epsilon_2(t)}{\rho_2} \leq \frac{\Delta}{4}$$

This further implies that,

$$\left\|\hat{\mathbf{U}}(t) - \mathbf{U}\right\|_{\infty,\infty} \leq \frac{\Delta}{2}$$

This guarantees that we select the optimal arm at time-step t. □

The following lemma we prove that each entry of the matrices $\hat{\mathbf{F}}(t)$ and $\hat{\mathbf{M}}_i(t)$ for all $i \in [l+1]$ are sampled sufficient number of times. Let $T_{sj}(t)$ denote the the number of samples obtained for the entry $\hat{\mathbf{F}}(t)_{sj}$. Similarly we define $N^{(i)}(t)_{sj}$ as the number of sampled for the enrty $\hat{\mathbf{M}}_i(t)_{sj}$.

Lemma 10. *Suppose* $\epsilon_t = \frac{(m+2m')\theta}{\beta t}$ *where* $\beta = \min_s \beta_s$. *Algorithm 1 ensures that,*

$$\mathbb{P}\left(T_{sj}(t) < \frac{\theta}{2}H_t\right) \leq \frac{1}{t^{\theta/12}}$$
$$\mathbb{P}\left(N^{(i)}(t)_{sj} < \frac{\theta}{2}H_t\right) \leq \frac{1}{t^{\theta/12}}$$

and where $H_n = \sum_{i=1}^{n}\frac{1}{i} \sim \log(n)$

Proof. Let S_t denote the random variable describing the context at time t. Let C_t denote the random variable denoting the the column of $\mathbf{G}(0)$ to be sampled provide $E(t) = 1$ and $H_t = 1$. Note that,

$$\begin{aligned}
\mathbb{E}\left[T_{sj}(t)\right] &\geq \sum_{l=1}^{t}\mathbb{P}\left(S_l = s, E(l) = 1, H_l = 1, C_l = j\right) \\
&\geq \sum_{l=1}^{t}\frac{\theta}{l} = \theta H_t
\end{aligned}$$

Now, a straight forward application of Chernoff-Hoeffding's inequality yields,

$$\begin{aligned}
\mathbb{P}\left(T_{sj}(t) < (1-\delta)\mathbb{E}\left[T_{sj}(t)\right]\right) &\leq \exp\left(-\frac{\delta^2}{3}\mathbb{E}\left[T_{sj}(t)\right]\right) \\
&\leq \exp\left(-\frac{\delta^2}{3}\theta H_t\right)
\end{aligned}$$

We can set $\delta = 1/2$ to get the required result. The same analysis works for $N^{(i)}(t)_{sj}$. The corresponding entry is sampled if $S_t = s_s(i)$. Let C'_t denote the column of $\mathbf{G}(i)$ to be sampled when $E(t) = 1$,$S_t = s_s(i)$ and $H_t = 0$.

$$\begin{aligned}
\mathbb{E}\left[N^{(i)}(t)_{sj}\right] &\geq \sum_{l=1}^{t}\mathbb{P}\left(E(t) = 1, S_t = s_s(i), H_t = 0, C'_l = j\right) \\
&\geq \sum_{l=1}^{t}\frac{\theta}{l} = \theta H_t
\end{aligned}$$

□

The same concentration inequality as before applies.

Lemma 11. *Under the conditions of Lemma 10 we have,*

$$\begin{aligned}
&\mathbb{P}\left(\left\|\hat{\mathbf{F}}(t) - \mathbf{F}\right\|_{\infty,\infty} > \epsilon_1(t)\right) \\
&\leq 4Lm'\exp\left(-\frac{\epsilon_1(t)^2}{2}\frac{\theta\log(t)}{2}\right) + \frac{2Lm'}{t^{\theta/12}}
\end{aligned}$$

Proof. The proof of this lemma is an application of Chernoff's bound to the samples observed. Note that $\mathbb{E}\left[\hat{\mathbf{F}}(t)\right] = \mathbf{F}$. We have,

$$\begin{aligned}
&\mathbb{P}\left(|\hat{\mathbf{F}}(t)_{sj} - \mathbf{F}_{sj}| > \epsilon_1(t)\right) \\
&\leq \mathbb{P}\left(|\hat{\mathbf{F}}(t)_{sj} - \mathbf{F}_{sj}| > \epsilon_1(t)\middle| T_{sj}(t) \geq \frac{\theta}{2}H_t\right) \\
&+ \mathbb{P}\left(T_{sj}(t) < \frac{\theta}{2}H_t\right) \\
&\leq 2e^{-\frac{\epsilon_1(t)^2}{2}\frac{\theta\log(t)}{2}} + \frac{1}{t^{\theta/12}}
\end{aligned}$$

where the last inequality if due to lemma 10. Now, we can apply an union bound over all $s \in [L]$ and $j \in [m]'$ to obtain the required result. □

Similarly we can bound the errors in estimating $\mathbf{M}_i$'s as in the lemma below.

Lemma 12. *Under the conditions of Lemma 10 we have,*

$$\mathbb{P}\left(\cup_{i\in[l+1]}\left\{\left\|\hat{\mathbf{M}}_i(t)-\mathbf{M}_i\right\|_{\infty,\infty}>\epsilon_2(t)\right\}\right)$$
$$\leq 4(K+1)m'\exp\left(-\frac{\epsilon_2(t)^2}{2}\frac{\theta\log(t)}{2}\right)+\frac{2(K+1)m'}{t^{\theta/12}}$$

Proof. The proof of this lemma is analogous to that of Lemma 11. We have the following chain,

$$\mathbb{P}\left(|\hat{\mathbf{M}}_i(t)_{sj}-\mathbf{M}_{\mathbf{i}sj}|>\epsilon_1(t)\right)$$
$$\leq \mathbb{P}\left(|\hat{\mathbf{M}}_i(t)_{sj}-\mathbf{M}_{\mathbf{i}sj}|>\epsilon_1(t)\middle|T_{sj}(t)\geq\frac{\theta}{2}H_t\right)$$
$$+\mathbb{P}\left(T_{sj}(t)<\frac{\theta}{2}H_t\right)$$
$$\leq 2e^{-\frac{\epsilon_2(t)^2}{2}\frac{\theta\log(t)}{2}}+\frac{1}{t^{\theta/12}}$$

We can apply union bound over all the entries of all the $l+1$ matrices to get the result. □

Now, we are at a position to prove our main theorem.

Proof of Theorem 8. We have $\epsilon_t=\frac{(m+2m')\theta}{\beta t}$ where we set,

$$\theta\geq 4\max\left(\frac{2m'((16+\Delta)\rho_2+32m)}{\Delta\rho_1\rho_2},\frac{15}{\rho_1(1-\lambda)}\right)^2 \tag{35}$$

By virtue of the ℓ_1-WStRIP property of $\mathbf{W}$, the set S is ρ_1-simplical with probability at least $1-\delta$. Similarly, by Lemma 7 all the sets $S(i)$ are *good* with probability at least $1-\delta$. In what follows, we will assume that the above high probability conditions hold. Note that according to Lemmas 11 and 12 we have,

$$\mathbb{P}\left(\left\|\hat{\mathbf{F}}(t)-\mathbf{F}\right\|_{\infty,\infty}>\frac{2}{\sqrt{\theta}}\right)$$
$$\leq\frac{4Lm'}{t}+o\left(\frac{1}{t^2}\right)$$
$$\mathbb{P}\left(\cup_{i\in[l+1]}\left\{\left\|\hat{\mathbf{M}}_i(t)-\mathbf{M}_i\right\|_{\infty,\infty}>\frac{2}{\sqrt{\theta}}\right\}\right)$$
$$\leq\frac{4(K+1)m'}{t}+o\left(\frac{1}{t^2}\right) \tag{36}$$

As $\mathbf{U}\in[0,1]^{L\times K}$ the regret till time T can be bounded as follows,

$$R(T)\leq\sum_{t=1}^{T}\mathbb{E}\left[\mathbb{1}\left\{E(t)=1\right\}\right]$$
$$+\sum_{t=1}^{T}\mathbb{E}\left[\mathbb{1}\left\{E(t)=0\right\}\right]\mathbb{P}\left(k(t)\neq k^*(s_t)\right) \tag{37}$$

By Lemma 9 we have that,

$$\mathbb{P}\left(k(t)\neq k^*(s_t)\right)\leq\mathbb{P}\left(\left\|\hat{\mathbf{F}}(t)-\mathbf{F}\right\|_{\infty,\infty}>\frac{2}{\sqrt{\theta}}\right)$$
$$+\mathbb{P}\left(\cup_{i\in[l+1]}\left\{\left\|\hat{\mathbf{M}}_i(t)-\mathbf{M}_i\right\|_{\infty,\infty}>\frac{2}{\sqrt{\theta}}\right\}\right)$$

We can combine this with (37) to get,

$$R(T)\leq\frac{\theta(m+2m')\log(T)}{\beta}+4(L+K+1)m'\log(T)+o(1)$$
$$=O\left(L\mathrm{poly}(m,m')\log(T)\right)$$

if we assume that $1/\beta=O(L)$. □

A.11 Lower Bound for α-consistent Policies

In this section we provide a problem dependent lower bound for the contextual bandit problem with *latent* contexts. The lower bound is established for a particular class of data-matrix $\mathbf{U}$ and for α-consistent policies. For, any $z_i\in\mathcal{Z}$ we define $\mathcal{C}(z_i)$ as,

$$\mathcal{C}(z_i):=\{s\in\mathcal{S}:\alpha_{si}\neq 0\}$$

Theorem 9. *Consider a problem instance $(\mathbf{U},\mathbf{A},\mathbf{W})$ such that $\beta_s=1/L$ for all $s\in\mathcal{S}$ and $|\mathcal{C}(z_i)|=L/m$ (assume that m divides L) for all $z_i\in\mathcal{Z}$. Further, we assume that $\mathcal{C}(z_i)\cap\mathcal{C}(z_j)=\emptyset$, for all $z_i\neq z_j$. Then the regret of any α-consistent policy is lower-bounded as follows,*

$$R(T)\geq(K-1)mD(\mathbf{U})\left((1-\alpha)(\log(T/2m)-\log(L/m))\right.$$
$$\left.-\log(4KC)\right)$$

for any $T>\tau$, where C,τ are universal constants independent of problem parameters and $D(\mathbf{U})$ is a constant that depends on the entries of $\mathbf{U}$ and is independent of L,K and m.

In order to prove Theorem 9 we introduce an inequality from the hypothesis testing literature.

Lemma 13 ([38])**.** *Consider two probability measures P and Q, both absolutely continuous with respect to a given measure. Then for any event $\mathcal{A}$ we have:*

$$P(\mathcal{A})+Q(\mathcal{A}^c)\geq\frac{1}{2}\exp\{-\min(\mathrm{KL}(P||Q),\mathrm{KL}(Q||P))\}$$

Proof of Theorem 9. Note that the conditions in the theorem imply that there are m distinct *latent* contexts and there are $L/m-1$ copies for each of them. For any $z_i\in\mathcal{Z}$ let us define $T(z_i)=\sum_{t=1}^{T}\mathbb{1}\left\{S_t\in\mathcal{C}(z_i)\right\}$. With some abuse of notation we also define $k^*(z_i)$ as the index of the optimal arm and $\Delta(z_i)$ as the gap between the optimal and second optimal arm for all contexts in $\mathcal{C}(z_i)$. By the assumptions in the theorem we have,

$$\mathbb{E}\left[T(z_i)\right] = \frac{T}{m}$$

Let E_i be the event $\left\{\frac{T}{2m} \leq T(z_i) \leq \frac{2T}{m}\right\}$. Let $E^c = \{\cup_{z_i \in \mathcal{Z}} E_i^c\}$. By a simple application of Chernoff bound we have,

$$\mathbb{P}\left(\cup_{z_i \in \mathcal{Z}} E_i^c\right) \leq 2me^{-T/12} = o\left(\frac{1}{T^2}\right)$$

Fix a $z_i \in \mathcal{Z}$ and let k be the index of an arm that is not optimal for any of the contexts that belong to $\mathcal{C}(z_i)$. Let us create another system with parameter $(\mathbf{U}', \boldsymbol{A}, \mathbf{W}')$ where we make the entry $W_{ik} = \lambda = \frac{U_{max}+1}{2}$ where $U_{max} = \max_{s,k} U_{sk}$, while everything else remains the same including the coefficients of the convex combinations relating the observed contexts to the *latent* contexts. Note that this implies that in the second system arm k is optimal for all $s \in \mathcal{C}(z_i)$. Let A be the event defined as follows,

$$A := \left\{ \sum_{\{t: S_t \in \mathcal{C}(z_i)\}} \mathbb{1}\left\{X_t = k\right\} \geq \frac{T(z_i)}{2} \right\}$$

Now, in the system with parameter $\mathbf{U}$ for any $s \in \mathcal{C}(z_i)$ we have,

$$\mathbb{E}\left[\sum_{\{t: S_t = s\}} \mathbb{1}\left\{X_t = k\right\}\right] \leq CT(s)^{\alpha}$$

if $T(s) \geq \tau$, since the policy in consideration is α-consistent. Here, τ, C are universal constants. By an application of Jensen's inequality we have,

$$\mathbb{E}\left[\sum_{\{t: S_t \in \mathcal{C}(z_i)\}} \mathbb{1}\left\{X_t = k\right\}\right] \leq C|\mathcal{C}(z_i)|^{1-\alpha} T(z_i)^{\alpha}$$

Let $\mathbb{P}_{\mathbf{U}}^T$ and $\mathbb{P}_{\mathbf{U}'}^T$ be the distributions corresponding to the chosen arms and rewards obtained for T plays for the two instances under a fixed α-consistent policy. Now we can apply Markov's inequality to conclude that,

$$\begin{aligned} \mathbb{P}_{\mathbf{U}}(A) &\leq \frac{2C|\mathcal{C}(z_i)|^{1-\alpha}}{T(z_i)^{1-\alpha}} \\ \mathbb{P}_{\mathbf{U}'}(A^c) &\leq \frac{2(K-1)C|\mathcal{C}(z_i)|^{1-\alpha}}{T(z_i)^{1-\alpha}} \end{aligned} \tag{38}$$

Now from Lemma 13 we have,

$$\begin{aligned} &\mathrm{KL}\left(\mathbb{P}_{\mathbf{U}}^T, \mathbb{P}_{\mathbf{U}'}^T\right) \\ &\geq (1-\alpha)\left(\log(T(z_i)) - \log(L/m)\right) - \log(4KC) \end{aligned}$$

Using standard methods from the bandit literature it can be shown that,

$$\mathrm{KL}\left(\mathbb{P}_{\mathbf{U}}^T, \mathbb{P}_{\mathbf{U}'}^T\right) = \sum_{s \in \mathcal{C}(z_i)} \sum_{\{t: S_t = s\}} \mathrm{KL}\left(U_{sk}, \lambda\right) \mathbb{E}_{\mathbf{U}}\left[\mathbb{1}\left\{X_t = k\right\}\right]$$

Let us define the regret incurred during the time-steps where $S_t \in \mathcal{C}(z_i)$ as $R(T(z_i))$. We can follow the same procedure for all the sub-optimal arms which yields the following bound,

$$\begin{aligned} &R(T(z_i)) \geq \Delta(z_i) \sum_{k \neq k^*(z_i)} \sum_{s \in \mathcal{C}(z_i)} \sum_{\{t: S_t = s\}} \mathbb{E}_{\mathbf{U}}\left[\{X_t = k\}\right] \\ &\geq \left(\operatorname*{argmin}_k \frac{(K-1)\Delta(z_i)}{\mathrm{KL}\left(U_{sk}, \lambda\right)}\right)\left((1-\alpha)\left(\log(T(z_i)) - \log(L/m)\right)\right. \\ &\left.- \log(4KC)\right) \end{aligned}$$

Let $D(\mathbf{U}) = \left(\operatorname{argmin}_{z_i,k} \frac{(K-1)\Delta(z_i)}{\mathrm{KL}(U_{sk},\lambda)}\right)$. Now, we have

$$\begin{aligned} &R(T) = \sum_{z \in \mathcal{Z}} \mathbb{E}\left[R(T(z_i))\right] \\ &\geq D(\mathbf{U})(K-1)\mathbb{E}\sum_{z \in \mathcal{Z}} \left((1-\alpha)\left(\log(T(z_i)) - \log(L/m)\right)\right. \\ &\left.- \log(4KC)\right) \end{aligned}$$

Now, using the fact that $T(z_i) \geq \frac{T}{2m}$ given E, we have

$$\begin{aligned} &R(T) = \sum_{z \in \mathcal{Z}} \mathbb{E}\left[R(T(z_i))\right] \\ &= \sum_{z \in \mathcal{Z}} \mathbb{E}\left[R(T(z_i))|E\right]\mathbb{P}(E) + \mathbb{E}\left[R(T(z_i))|E^c\right]\mathbb{P}(E^c) \\ &\geq D(\mathbf{U})(K-1)m\left((1-\alpha)\left(\log(T/2m) - \log(L/m)\right) -\right. \\ &\left.\log(4KC)\right) + o(1) \end{aligned}$$

$\square$