Cluster Agnostic Network Lasso Bandits

Sofien Dhouib[*]
Technical University of Darmstadt

Steven Bilaj[*]
University of Tübingen

Behzad Nourani-Koliji
University of Tübingen

Setareh Maghsudi
Ruhr University Bochum

Abstract

We consider a multi-task contextual bandit setting, where the learner is given a graph encoding relations between the bandit tasks. The tasks' preference vectors are assumed to be piecewise constant over the graph, forming clusters. At every round, we estimate the preference vectors by solving an online network lasso problem with a suitably chosen, time-dependent regularization parameter. We establish a novel oracle inequality relying on a convenient restricted eigenvalue assumption. Our theoretical findings highlight the importance of dense intra-cluster connections and sparse inter-cluster ones. That results in a sublinear regret bound significantly lower than its counterpart in the independent task learning setting. Finally, we support our theoretical findings by experimental evaluation against graph bandit multi-task learning and online clustering of bandits algorithms.

1 Introduction

Online commercial websites aim to recommend their products to their customers properly, and the performance of these recommendations depends on the knowledge of users' preferences. Unlike traditional collaborative-filtering-based methods [Su and Khoshgoftaar, 2009], such knowledge is initially unavailable. Therefore, the online recommender systems need to recommend various items to the users and observe their ratings to *explore* their preferences. At the same time, the recommender system should be able to recommend items that attract users' attention and receive high ratings by *exploiting* the learned knowledge. The contextual bandit frameworks [Li et al., 2010] have been popularly used to formalize and address this exploration-exploitation trade-off.

However, the classical form of contextual bandits [Li et al., 2010, Chu et al., 2011, Abbasi-Yadkori et al., 2011] ignores the availability of social networks amongst users and solves the problem for each user separately. Consequently, such algorithms have some drawbacks when applied to problems with a large number of users. First, such a large number hinders their computational efficiency. Second, the partial feedback of the bandit settings exposes the algorithms to having weak estimations and impairing their decision-making ability [Yang et al., 2020]. Consequently, to improve bandit algorithms' performance for large-scale applications, structural assumptions that link the different users are usually integrated within bandit algorithms [Cesa-Bianchi et al., 2013, Gentile et al., 2014, Li et al., 2019, Herbster et al., 2021].

Cesa-Bianchi et al. [2013] and Yang et al. [2020] use the prior knowledge of social networks into their contextual bandit algorithms. Both papers propose UCB-style algorithms and exhibit the importance of using the social network graph to achieve lower regrets using Laplacian regularization. The latter regularization promotes smoothness among the preference vectors of users, allowing the transfer of the collected information between them. However, the Laplacian regularization does not account for

[*]Equal contribution

18th European Workshop on Reinforcement Learning (EWRL 2025).

the smoothness heterogeneity introduced by a piecewise constant behavior over the graph [Wang et al., 2016]. On the other hand, algorithms of online clustering of bandits [Gentile et al., 2014, Li et al., 2019] tackle such a piecewise constant behavior by explicitly estimating user clusters. However, their clustering can cause overconfidence in the constructed clusters, potentially leading to error accumulation.

In this paper, we assume access to a graph encoding relations between bandit tasks, and that the task parameter vectors are piecewise constant over the graph. We propose an algorithm that integrates the prior knowledge of the piecewise constant structure to update tasks rather than finding the clusters explicitly. That way, we mitigate the limitations mentioned above: the piecewise constant smoothness is naturally integrated into our regularizer, and we do not estimate the clusters so our algorithm does not suffer from overconfidence drawbacks.

More precisely, we provide the following contributions

- We analyze an instance of the Network Lasso problem [Hallac et al., 2015], estimating every vertex's preference vector using data generated during the interaction between users and the bandit. We provide the first oracle inequality in this setting and link it to fundamental quantities characterizing the relation between the graph and the true preference vectors of the users. Our result relies on our novel restricted eigenvalue (RE) condition, which we assume for our setting. This result is of independent interest and can be applied to i.i.d. data as a special case.
- We prove that the empirical multi-task Gram matrix of the data inherits the RE condition from its true counterpart. Both this result and the previous one depend on the sparsity of inter-cluster connections and the density of intra-cluster ones.
- We provide a regret upper bound for our setting. Our bound highlights the advantage of our algorithm in high dimensional settings, and for large graphs.
- We support our theoretical findings by extensive numerical experiments on simulated data that prove the advantage of our algorithm over other related approaches.

The rest of the paper is organized as follows. Section 2 discusses the relation of our work to the literature. We formulate our problem and state some of our assumptions in Section 3, then present our bandit algorithm in Section 4. We analyze the problem theoretically in Section 5 and demonstrate its practical interest experimentally in Section 6.

2 Related work

Lasso contextual bandits. To address the high dimensional setting for linear bandits, several multi-armed bandit papers solve a LASSO [Tibshirani, 1996] problem under different assumptions [Bastani and Bayati, 2019, Kim and Paik, 2019, Oh et al., 2021, Ariu et al., 2022]. They all rely on a previously established compatibility or RE condition [Bühlmann and van de Geer, 2011], that they adapt to the non-i.i.d case resulting from the context selection procedure across rounds. Such assumptions were also used in the multi-task setting by Cella and Pontil [2021] with a Group Lasso regularization [Yuan and Lin, 2006], and to impose a low-rank structure on the task preference vectors in Cella et al. [2023]. In our case, we establish a novel oracle inequality, rather than only generalize an existing one to the non-i.i.d setting, with a newly introduced RE assumption, which can be of independent interest.

Clustering of bandits. Gentile et al. [2014] introduced sequential clustering of bandits with the CLUB algorithm. The latter starts with a fully connected graph, and then an iterative graph learning process is performed, where edges between users are deleted if their preference vectors are significantly different. As a result, any connected component is seen as a cluster and only one recommendation per cluster is developed. The SCLUB algorithm of Li et al. [2019] generalizes CLUB via including merging operations in addition to splitting. In contrast to these approaches, Nguyen and Lauw [2014] groups users via K-means clustering, and Cheng et al. [2023] rely on hedonic games for online clustering of bandits. Furthermore, Yang and Toni [2018] make use of community detection techniques on graphs to find user clusters. Gentile et al. [2017] study the clustering of the contextual bandit problem where their proposed algorithm, named CAB, adaptively matches user preferences in the face of constantly evolving items. Our work fundamentally differs from the previous ones on two aspects. First, we assume access to a graph encoding relations between users, which is more informative than a complete graph. Second, we do not keep track of a model for each cluster, but rather we integrate a prior over the graph via a graph total variation regularizer that enforces a piecewise constant behavior for the estimated preference vectors.

Multi-task learning. Several contributions assume that the bandit tasks share some underlying structure. In Cella and Pontil [2021], task preference vectors are assumed to be sparse and to share their sparsity support, implying that they lie in a low-dimensional subspace with dimensions aligning with the canonical basis vectors. This idea is further generalized in Cella et al. [2023], where the tasks are assumed to be confined to an arbitrary unknown low-dimensional subspace. That work improves upon Hu et al. [2021] by not requiring the knowledge of the small dimension of the task space. It can be considered to solve our problem if the number of clusters is smaller than the dimension, resulting in a low-rank structure. However, our work does not rely on any assumption between the number of clusters and the dimension. The underlying structure linking tasks can also be a graph encoding relations between them [Cesa-Bianchi et al., 2013, Yang and Toni, 2018], which is our case. However, while they assume smoothness as a prior, we assume piecewise constant behavior.

Homophily and modularity in social networks Given the large number of users on social networks, one may be able to learn their preferences more quickly by leveraging the similarities between them. This idea relies on the notion of *homophily* in social networks [McPherson et al., 2001, Easley et al., 2010]. In modelling social networks, users' preferences relationships are encoded in a graph, where neighboring nodes are users with similar preferences. This graph can be known *a priori* or it can be inferred from previously collected feedback [Dong et al., 2019]. Exploiting this information and integrating them into bandit algorithms can lead to a significant increase in performance Yang et al. [2020]. Indeed, the knowledge of user relations allows the algorithm to tackle the data sparsity issue that is inherent to bandit settings. Another fundamental point that can be used to integrate information from social networks is that, social networks show large *modularity* measures [Newman, 2006, Borge-Holthoefer et al., 2011]. This implies that we have high density of edges within clusters and low density of edges between clusters. As a result, users can be clustered based on the graph topology and a preference vector can be learned for each cluster, substantially reducing the dimensionality of the problem. In other words, discovering the clustering structure of users can reduce the computational burden of large social networks. Consequently, there have been attempts in exploiting the clustered structures of social networks in bandit algorithms [Gentile et al., 2014, Nguyen and Lauw, 2014, Yang and Toni, 2018, Li et al., 2019, Nourani-Koliji et al., 2023, Cheng et al., 2023].

3 Problem setting

We consider a linear bandit setting, with a finite number of tasks representing users in a recommendation system for example. For each task the agent has to choose among K arms, each associated to a d-dimensional context vector. All interactions over a horizon of T time steps. We further assume that we have access to an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with vertex set $\mathcal{V}$ representing the tasks and edge set $\mathcal{E}$ encoding the relationships between them. We identify the vertex set $\mathcal{V}$ with the set of vertex indices $[|\mathcal{V}|]$. Thus, we consider $\mathcal{E}$ to be a subset of $\mathcal{V}^2$, where every edge $(m, n) \in \mathcal{E}$ has weight $w_{mn} > 0$, with $m < n$. The tasks' preference vectors are denoted by $\{\boldsymbol{\theta}_m\}_{m \in \mathcal{V}} \subset \mathbb{R}^d$ verifying $\|\boldsymbol{\theta}_m\| \leq 1 \; \forall m \in \mathcal{V}$, which we concatenate as row vectors into matrix $\boldsymbol{\Theta} \in \mathbb{R}^{|\mathcal{V}| \times d}$. The latter represents a graph vector signal, assumed to be piecewise constant over $\mathcal{G}$.

At a round $t \in \mathbb{N}^\star$, a user $m(t) \in \mathcal{V}$ is selected uniformly at random and served an arm with context vector $\boldsymbol{x}(t)$ from a finite action set $\mathcal{A}(t) \subset \mathbb{R}^d$ with size K, depending on their estimated preference vector $\hat{\boldsymbol{\theta}}_{m(t)}(t) \in \mathbb{R}^d$. We assume the expected reward to be linear, with an additive σ-sub-Gaussian noise conditionally on the past. Formally, denoting by $\mathcal{F}_0$ the trivial sigma-algebra, and for all $t \geq 1$, by $\mathcal{F}_t$ the sigma-algebra generated by history set $\{m(1), \boldsymbol{x}(1), y(1), \cdots, m(t), \boldsymbol{x}(t), y(t), m(t+1)\}$, the received reward $y(t)$ is given by $y(t) = \langle \boldsymbol{\theta}_{m(t)}(t), \boldsymbol{x}(t) \rangle + \eta(t)$, where $\eta(t)$ is $\mathcal{F}_t-$measurable and $\forall t \geq 1, \forall s \in \mathbb{R}$,

$$\mathbb{E}\left[\eta(t) | \mathcal{F}_{t-1}\right] = 0, \quad \mathbb{E}\left[\exp(s\eta(t)) | \mathcal{F}_{t-1}\right] \leq \exp\left(\frac{1}{2}\sigma^2 s^2\right). \tag{1}$$

The performance of our policy is assessed by the expected regret over the T interaction rounds for all tasks:

$$\mathcal{R}(T) = \mathbb{E}\left[\sum_{t=1}^{T} \max_{\tilde{\boldsymbol{x}} \in \mathcal{A}(t)} \left\langle \boldsymbol{\theta}_{m(t)}, \tilde{\boldsymbol{x}} \right\rangle - \left\langle \boldsymbol{\theta}_{m(t)}, \boldsymbol{x}(t) \right\rangle \right]. \tag{2}$$

The Optimization problem in Equation (4) is an instance of the Network Lasso [Hallac et al., 2015]. Several instances of the same type were studied by Jung et al. [2018], Jung and Vesselinova [2019], Jung [2020], He et al. [2019]. The objective is characterized by its second term which, while being just the Laplacian regularization without squaring the norms, promotes a piecewise constant behavior rather than smoothness. For real-valued signals ($d = 1$), this regularization has been extensively studied for image and graph signal denoising, for the problem of trend filtering on graphs [Wang et al., 2016]. According to Wang et al. [2016], that regularization better adapts to the heterogeneity of smoothness of the signal and induces a cluster structure in the data: similar users will not only have similar models but the same model, which offers a compression of the overall model over the graph. Note that our setting is cluster agnostic; our algorithm does not aim to learn the cluster structure explicitly but to exploit it implicitly using the total variation semi-norm as regularization. The strength of the latter is controlled via a time-dependent regularization coefficient $\alpha(t)$, which we will express later in the analysis.

We formalize our assumption on the context generation as follows.

Assumption 1 (i.i.d action sets)**.** *Context sets $\{\mathcal{A}(t)\}_{t=1}^{T}$ are generated i.i.d. from a distribution p over $\mathbb{R}^{K\times d}$, such that $\|\boldsymbol{x}\| \leq 1 \forall\, \boldsymbol{x} \in \mathcal{A}(t)\ \forall t \geq 1$.*

In addition to the i.i.d assumption, we assume more regularity as follows.

Assumption 2 (Relaxed symmetry and balanced covariance)**.** *There exists a constant $\nu \geq 1$ such that for all $\boldsymbol{X} \in \mathbb{R}^{K\times d}$, $p(-\boldsymbol{X}) \leq \nu p(\boldsymbol{X})$. Furthermore, there exists $\omega > 0$, such that for any permutation $(a_1, \cdots, a_K)$ of $[K]$, for any $i \in \{2, \cdots, K-1\}$, $\boldsymbol{w} \in \mathbb{R}^d$, we have*

$$\mathbb{E}\left[\boldsymbol{x}_{a_i}\boldsymbol{x}_{a_i}^\top [\boldsymbol{w}^\top \boldsymbol{x}_{a_1} < \cdots < \boldsymbol{w}^\top \boldsymbol{x}_{a_K}]\right] \preccurlyeq \omega \mathbb{E}\left[(\boldsymbol{x}_{a_1}\boldsymbol{x}_{a_1}^\top + \boldsymbol{x}_{a_K}\boldsymbol{x}_{a_K}^\top)[\boldsymbol{w}^\top \boldsymbol{x}_{a_1} < \cdots < \boldsymbol{w}^\top \boldsymbol{x}_{a_K}]\right],$$

where $\boldsymbol{M} \preccurlyeq \boldsymbol{N}$ means that $\boldsymbol{N} - \boldsymbol{M}$ is a PSD matrix.

This assumption was introduced in Oh et al. [2021], and has already been used in a multi-task setting by Cella et al. [2023]. Parameter ν controls the skewness, as $\nu = 1$ corresponds to a symmetric distribution. ω decreases with increasing positive correlation between arms. It verifies $\omega = O(1)$ for multi-variate Gaussians and uniform distributions over the unit sphere [Oh et al., 2021]. The piecewise constant behavior of the graph signal $\boldsymbol{\Theta}$ is formalized in the next assumption.

Assumption 3 (Piecewise constant signal)**.** *There exists a partition $\mathcal{P}$ of $\mathcal{V}$, such that for any cluster $\mathcal{C} \in \mathcal{P}$, signal $\boldsymbol{\Theta}$ is constant on $\mathcal{C}$, and the graph obtained by taking the vertices in $\mathcal{C}$ and the edges linking them is connected.*

Assumption 3 basically states that the true preference vectors are clustered and that the given graph induces the cluster structure. It is required for our approach to be beneficial, as we will detail in the analysis section. For the sake of clarity, we defer the statement of other technical assumptions to Section 5.

4 Algorithm

Our policy in Algorithm 1 follows a greedy arm selection rule in a multi-task setting, in the same vein as those presented in Oh et al. [2021], Cella et al. [2023]. Indeed, as pointed out in Oh et al. [2021], exploration is implicitly incorporated into regularization parameter $\alpha(t)$'s time dependence. It has the following expression

$$\begin{aligned}
\alpha(t) &:= \frac{\alpha_0 \sigma}{t}\sqrt{t + \alpha_1(t) + \alpha_2(t)}, \qquad (3)\\
\alpha_1(t) &:= \sqrt{2\sum_{m\in\mathcal{V}} |\mathcal{T}_m(t)|^2 \log\frac{1}{\delta(t)}},\\
\alpha_2(t) &:= 2\max_{m\in\mathcal{V}} |\mathcal{T}_m(t)| \log\frac{1}{\delta(t)},
\end{aligned}$$

where $\alpha_0 > 0$. The set of time steps a task m has been selected up to time t is denoted by $\mathcal{T}_m(t)$. At the end of a round t, all preference vectors are updated into a new estimation $\hat{\mathbf{\Theta}}(t)$ while leveraging the structure of graph $\mathcal{G}$, formally by solving the following network lasso optimization problem:

$$\hat{\mathbf{\Theta}}(t) = \underset{\tilde{\mathbf{\Theta}} \in \mathbb{R}^{|\mathcal{V}| \times d}}{\arg\min} \frac{1}{2t} \sum_{\tau=1}^{t} \left(\left\langle \tilde{\boldsymbol{\theta}}_{m(\tau)}, \boldsymbol{x}(\tau) \right\rangle - y(\tau) \right)^2 + \alpha(t) \sum_{(m,n) \in \mathcal{E}} w_{mn} \left\| \tilde{\boldsymbol{\theta}}_m - \tilde{\boldsymbol{\theta}}_n \right\|, \quad (4)$$

where $\|\cdot\|$ denotes the Euclidean norm for vectors. At each time step the network Lasso problem is solved via the primal-dual algorithm [Jung, 2020].

Algorithm 1: Network Lasso Policy

Input: $T, \alpha_0 > 0, \mathcal{G}, \delta$
Initialization: $\hat{\mathbf{\Theta}}(0) = \mathbf{0} \in \mathbb{R}^{|\mathcal{V}| \times d}$
for $t \in \{1, ..., T\}$ **do**
  Draw a user $m(t) \in \mathcal{V}$ uniformly at random
  Observe context set $\mathcal{A}(t)$
  Select $\boldsymbol{x}(t) \in \arg\max_{\tilde{\boldsymbol{x}} \in \mathcal{A}(t)} \left\langle \hat{\boldsymbol{\theta}}_{m(t-1)}, \tilde{\boldsymbol{x}} \right\rangle$
  Receive payoff $y(t)$
  Update $\alpha(t)$ via equation 3
  Update $\hat{\mathbf{\Theta}}(t)$ via equation 4

5 Analysis

This section provides the main steps of the analysis. One of the paper's contribution lies in finding an oracle inequality of the network lasso problem given a restricted eigenvalue condition holding for the true multi-task Gram matrix. In this regard, the next major challenge and contribution is to show that the empirical multi-task Gram matrix, estimated in the algorithm, satisfies the restricted eigenvalue condition. We start by proving an oracle inequality for the estimation error of $\mathbf{\Theta}$. Then, we prove that the latter assumption holds with high probability given that the true multi-task Gram matrix satisfies it. We end this section by establishing a regret bound for our algorithm.

5.1 Notation and technical assumptions

We provide additional notations required for the analysis. We denote by $\partial\mathcal{P}$ the set of all edges in $\mathcal{E}$ connecting vertices from different clusters from partition $\mathcal{P}$ (Assumption 3), and we call it the boundary of $\mathcal{P}$. Thus, $\partial\mathcal{P}^c$, the complementary set of $\partial\mathcal{P}$, is formed by edges connecting vertices of the same cluster. The total weight of the boundary, *i.e.*the sum of its edges' weights, is referred to as $w(\partial\mathcal{P})$. Given a signal $\boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}| \times d}$, we denote by $\overline{\boldsymbol{Z}}_{\mathcal{P}}$ the signal obtained by setting row vectors of $\boldsymbol{Z}$ to their mean-per-cluster value w.r.t. $\mathcal{P}$. For any edge subset $I \subseteq \mathcal{E}$, we denote the following norms: $\|\cdot\|_F$ as the Frobenius norm and $\|\mathbf{\Theta}\|_I := \sum_{(m,n) \in I} w_{mn} \|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\|$ as the total variation semi-norm of $\mathbf{\Theta} \in \mathbb{R}^{|\mathcal{V}| \times d}$ over I. Thus, the regularization term of Problem Equation (4) is equal to $\|\mathbf{\Theta}\|_{\mathcal{E}}$. Also, we define the incidence matrix $\boldsymbol{B}_I \subset \mathbb{R}^{|\mathcal{E}| \times |\mathcal{V}|}$restricted to $I \subseteq \mathcal{E}$ to be null except at rows with index $i \in I$ corresponding to edge (m, n), where it equals $w_{mn}(\boldsymbol{e}_m - \boldsymbol{e}_n)$, where $\boldsymbol{e}_m$ is the m^{th} canonical basis vector of $\mathbb{R}^{|\mathcal{V}|}$. We define $\boldsymbol{A}_{\mathcal{V}}(t) := \operatorname{diag}\left(\boldsymbol{X}_1(t)^\top \boldsymbol{X}_1(t), \ldots, \boldsymbol{X}_{|\mathcal{V}|}(t)^\top \boldsymbol{X}_{|\mathcal{V}|}(t)\right) \in \mathbb{R}^{d|\mathcal{V}| \times d|\mathcal{V}|}$, and subsequently the empirical multi-task Gram matrix up to time step t is given by $\frac{1}{t}\boldsymbol{A}_{\mathcal{V}}(t)$. The following definition introduces quantities related to the clusters defined by partition $\mathcal{P}$, with crucial roles that we will elucidate throughout the analysis.

Definition 1 (Cluster content constants). *Let $\mathcal{C} \in \mathcal{P}$ be a cluster.*

- *We denote by $\partial_v\mathcal{C}$ the vertices of $\mathcal{C}$ that are connected to its complementary. We define the inner isoperimetric ratio of $\mathcal{C}$ as $\iota_{\mathcal{G}}(\mathcal{C}) := \frac{|\partial_v \mathcal{C}|}{|\mathcal{C}|}$.*
- *By abuse of notation, we denote as $\boldsymbol{B}_{\mathcal{C}}$ the incidence matrix restricted to edges linking vertices of $\mathcal{C}$, its associated Laplacian matrix by $\boldsymbol{L}_{\mathcal{C}} := \boldsymbol{B}_{\mathcal{C}}^\top \boldsymbol{B}_{\mathcal{C}}$, and its pseudo-inverse by $\boldsymbol{L}_{\mathcal{C}}^\dagger$. The*

topological centrality index of node $m \in \mathcal{C}$ w.r.t $\mathcal{C}$ is equal to $(\boldsymbol{L}_{\mathcal{C}}^{\dagger})_{mm}^{-1}$. We define the topological centrality index of $\mathcal{C}$ by $c_{\mathcal{G}}(\mathcal{C}) := \min_{m \in \mathcal{C}} (\boldsymbol{L}_{\mathcal{C}}^{\dagger})_{mm}^{-1}$.

The inner isoperimetric ratio of a cluster measures how many "interior" nodes a cluster contains, in the sense that they are not connected to its complementary. It is at most equal to the isoperimetric ratio for weightless graphs as the size of the inner boundary is at most equal to that of the edge boundary, the latter being connected to the algebraic connectivity via the Cheeger inequality [Cheeger, 1970].

The topological centrality index measures the overall connectedness of a vertex in a network and indicates how robust a node is to edge failures [Ranjan and Zhang, 2013]. Also, it can be tied to electricity spreading in a network according to Van Mieghem et al. [2017]. We refer the interested reader to the two previously mentioned works for a detailed account of the properties of the topological centrality index. In the appendix, we show that for binary weights graphs the minimum topological centrality index is at least equal to the algebraic connectivity theoretically and experimentally, where we showcase that the difference between the two can be significant.

Remark 1. *Both the topological centrality index and inner isoperimetric ratio are key parameters of the cluster structure and the graph. They determine the 'quality' of the given graph. An optimal graph and cluster structure yield many intra-cluster connections and few inter cluster connections i.e. a high topological centrality index and low inner isoperimetric ratio for any cluster. This will later be highlighted in the oracle inequality and the regret bound.*

To proceed, we will need the following definition that introduces several notations to reduce the clutter.

Definition 2 (Restricted Eigenvalue (RE) condition and norm)**.** *A PSD matrix $\boldsymbol{M} \in \mathbb{R}^{d|\mathcal{V}| \times d|\mathcal{V}|}$ verifies the RE condition with constants $\kappa \geq 1$, $\psi > 0$ and $\phi > 0$ if*

$$\phi^2 \|\boldsymbol{Z}\|_{\mathrm{RE}}^2 \leq \mathrm{vec}(\boldsymbol{Z}^\top)^\top \boldsymbol{M} \,\mathrm{vec}(\boldsymbol{Z}^\top) \quad \forall \boldsymbol{Z} \in \mathcal{S}, \tag{5}$$

where $\mathcal{S}$ is the cone defined by:

$$\mathcal{S} := \left\{ \boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}| \times d}; a_1\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right) \|\boldsymbol{Z}\|_{\partial\mathcal{P}^c} \leq a_2\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right) \left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F \right\},$$

$$a_1(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := 1 - \frac{\frac{1}{\alpha_0} + 2\kappa w(\partial\mathcal{P})}{\min\limits_{\mathcal{C} \in \mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}}, \quad a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := \frac{1}{\alpha_0} + \sqrt{2}\kappa w(\partial\mathcal{P}) \max_{\mathcal{C} \in \mathcal{P}} \sqrt{\iota_{\mathcal{G}}(\mathcal{C})},$$

and the RE semi-norm is defined by $\|\boldsymbol{Z}\|_{\mathrm{RE}} := \left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F$.

For the rest of the paper, when we use a_1 and a_2 without arguments, we set $\alpha_0 = \frac{1}{w(\partial\mathcal{P})\psi}$ in order to reduce clutter. For our main results, we cover the case of $\kappa \geq 1$ but treat the more general case $\kappa > 0$ in the proofs in the supplementary material. For such a simplification to be valid, we need to assume that the graph satisfies $\min\limits_{\mathcal{C} \in \mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} > 2w(\partial\mathcal{P})$, such that we can always find constants $\kappa \geq 1$ and $\psi > 0$ that guarantee $a_1 > 0$

To explain the RE condition, if we had $\mathcal{S} = \mathbb{R}^{|\mathcal{V}| \times d}$ and $\|\cdot\|_{\mathrm{RE}} = \|\cdot\|_F$, then $\boldsymbol{M}$ would be invertible with minimum eigenvalue at least ϕ^2. In comparison, our requirement is weaker since it only needs to hold for signals $\boldsymbol{Z} \in \mathcal{S}$ and for the $\|\cdot\|_{\mathrm{RE}}$ semi-norm. It has the same form as the compatibility assumption for the Lasso problem in [Bühlmann and van de Geer, 2011, Oh et al., 2021] or the restricted strong convexity assumption [Cella et al., 2023].

We further make the following assumption on the true multi-task Gram matrix:

Assumption 4 (RE condition for the true multi-task Gram matrix)**.** *For $k \in [K]$, let $\boldsymbol{\Sigma}_k := \mathbb{E}\left[\boldsymbol{x}_k \boldsymbol{x}_k^\top\right]$ be the Gram matrix of the k^{th} context vector's marginal distribution, let $\boldsymbol{\Sigma}_{\mathcal{V}}$ be the true multi-task Gram matrix of the context vector generating distribution, given by*

$$\boldsymbol{\Sigma}_{\mathcal{V}} := \boldsymbol{I}_{|\mathcal{V}|} \otimes \overline{\boldsymbol{\Sigma}}, \quad \text{where} \quad \overline{\boldsymbol{\Sigma}} = \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{\Sigma}_k. \tag{6}$$

We assume that $\boldsymbol{\Sigma}_{\mathcal{V}}$ verifies RE condition (Definition 2) with some problem dependent constants $\kappa \in \left[1, \frac{1}{2w(\partial\mathcal{P})} \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}\right)$, $\psi \in \left(0, \frac{1}{w(\partial\mathcal{P})} \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} - 2\right)$ and $\phi > 0$.

This assumption is common to several Lasso-like bandit problems [Oh et al., 2021, Ariu et al., 2022, Cella et al., 2023].

We will later show that it can be transferred to the empirical multi-task Gram matrix.

We provide further intuition on the constant ϕ within the RE condition. We can show that ϕ has an upper bound:

Proposition 1 (On the RE constant ϕ)**.** *Let $M_i \in \mathbb{R}^{d\times d}$ be the true multi-task gram matrix of user i. Assume $\kappa \geq 1$. Then the constant ϕ of the RE condition can be upper bounded as:*

$$\phi \leq \sqrt{\lambda_{\min}\left(\frac{\sum_{i\in\mathcal{C}} \boldsymbol{M}_i}{|\mathcal{C}|}\right)},$$

where $\lambda_{\min}(\cdot)$ yields the minimum eigenvalue of a given matrix.

Since the true multi-task gram matrix per cluster is always invertible, we always have a non-null minimal eigenvalue.

Remark 2. *The minimal eigenvalue in Proposition 1 could be further bounded using the trace of the covariances i.e. the sum of all the eigenvalues over the dimension. This would result into an upper bound of $\phi^2 \leq \frac{1}{d}$.*

5.2 Oracle inequality

This section is dedicated to provide a bound on the estimation error of the Network Lasso problem given in Equation (4) at a particular step t of Algorithm 1.We assume fixed design, meaning that the context vectors are given and fixed, and we are not concerned by their randomness (due to the context generating distribution), nor by the randomness of their number for each user (due to random selection at each time step).

For a time step t, we deliver the oracle inequality controlling the deviation between the estimated preference vectors $\hat{\boldsymbol{\Theta}}(t)$ and the true ones $\boldsymbol{\Theta}$.

Theorem 1 (Oracle inequality)**.** *Assume that the RE assumption holds for the empirical multi-task Gram matrix $\dfrac{1}{t}\boldsymbol{A}_{\mathcal{V}}(t)$ with constants $\kappa \in \left[1, \frac{1}{2w(\partial\mathcal{P})} \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}\right)$, $\psi \in \left(0, \frac{1}{w(\partial\mathcal{P})} \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} - 2\right)$ and $\phi > 0$. Suppose that $\max_{m\in\mathcal{V}} |\mathcal{T}_m(t)| \leq bt$ for some $b > 0$ and $\alpha_0 \geq \frac{1}{\psi w(\partial\mathcal{P})}$. Then, with a probability at least $1-\delta(t)$, we have*

$$\left\|\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}(t)\right\|_F \leq 2\frac{\sigma\alpha_0}{\phi^2\sqrt{t}} f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \sqrt{1 + 2b\sqrt{|\mathcal{V}| \log\frac{1}{\delta(t)}} + 2b\log\frac{1}{\delta(t)}},$$

where

$$f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\left(\frac{a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)}{a_1(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}} + 1\right).$$

Due to the expressions of $a_1(\boldsymbol{\Theta}, \mathcal{G}, \alpha_0)$ and $a_2(\boldsymbol{\Theta}, \mathcal{G}, \alpha_0)$, the bound significantly decreases with the products $w(\partial\mathcal{P}) \min_{\mathcal{C}\in\mathcal{P}} \sqrt{\iota(\mathcal{C})}$ and $w(\partial\mathcal{P}) \max_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-\frac{1}{2}}$, which are small enough for dense intra-cluster edge links and sparse inter-cluster ones. The bound on the oracle inequality clearly grows with κ and ψ, thus it is most beneficial if κ is close to 1 and ψ close to zero.

5.3 RE condition for the empirical multi-task Gram matrix

To establish the oracle inequality, we assumed that the RE condition holds for the empirical multi-task Gram matrix. In this section, we prove that this holds with high probability. To this end, we use the same strategy as in Oh et al. [2021], Cella et al. [2023]. We prove that on the one hand, the empirical multi-task Gram matrix inherits the RE condition from its adapted counterpart since it concentrates around it. On the other hand, we show that the adapted Gram matrix verifies the RE condition due to Assumption 1, 2 and 4.

Theorem 2 (RE condition holding for the empirical multi-task Gram matrix)**.** *Under assumptions 2 and 4, let $t \geq 1$, and let κ, ϕ be the constants from Assumption 4. Assume that $\max_{m \in \mathcal{V}} |\mathcal{T}_m(t)| \leq bt$. Then, for any $\gamma \in \left(0, \left(1 + \frac{a_2}{a_1}\right)^{-2}\right)$, the empirical multi-task Gram matrix $\frac{1}{t}\boldsymbol{A}_{\mathcal{V}}(t)$ verifies the RE condition with constants κ, ψ and $\hat{\phi}$, where*

$$\hat{\phi} = \tilde{\phi}\sqrt{1 - \gamma\left(1 + \frac{a_2}{a_1}\right)^2}, \tag{7}$$

with a probability at least equal to $1 - 6d|\mathcal{V}| \exp\left(\dfrac{-3\gamma^2\tilde{\phi}^4(\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}_{\mathcal{G}}(\mathcal{C})^2)t}{6b + 2\sqrt{2}\gamma\tilde{\phi}^2}\right)$, where $\tilde{\phi} := \dfrac{\phi}{\sqrt{2\nu\omega}}$ and $\tilde{c}_{\mathcal{G}}(\mathcal{C}) := c_{\mathcal{G}}(\mathcal{C}) \wedge |\mathcal{C}| \quad \forall \mathcal{C} \in \mathcal{P}$.

The proof follows a similar approach as in Oh et al. [2021], Cella et al. [2023]; we prove that the RE condition transfers from the true multi-task Gram matrix to its adapted counterpart $\boldsymbol{V}_{\mathcal{V}}(t)$, defined as follows:

$$\boldsymbol{V}_{\mathcal{V}}(t) = \operatorname{diag}\left(\boldsymbol{V}_1(t), \cdots, \boldsymbol{V}_{|\mathcal{V}|}(t)\right), \tag{8}$$

where

$$\boldsymbol{V}_m(t) = \frac{1}{t}\sum_{\tau \in \mathcal{T}_m(t)} \mathbb{E}\left[\boldsymbol{x}(\tau)\boldsymbol{x}(\tau)^\top | \mathcal{F}_{\tau-1}\right]. \tag{9}$$

This transfer relies on the work of Oh et al. [2021, lemma 10]. The other step of the proof is showing that the empirical multi-task Gram matrix and $\boldsymbol{V}_{\mathcal{V}}(t)$ become close to each other with high probability after sufficiently many time steps, in the sense of a matrix norm induced by the RE semi-norm and the restriction to set $\mathcal{S}$ (Definition 2). The bound showcases a dependence on $\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C}) \wedge |\mathcal{C}|$, which is of the same order as $|\mathcal{C}|$ for a fully connected cluster with vertices $\mathcal{C}$. It is also clear that the probability of satisfying the RE condition increases with a higher minimum centrality of a cluster.

5.4 Regret bound

To bound the regret, we bound the expected instantaneous regret for each round $t \geq 1$. This bound relies on the oracle inequality holding and the RE condition being satisfied for the empirical Gram matrix, both with high probability. Thanks to Theorem 1 and Theorem 2, these two conditions are ensured.

Theorem 3. *Let the mean horizon per node be $\overline{T} = \frac{T}{|\mathcal{V}|}$. Under assumptions 1 to 4, the expected regret of the Network Lasso Bandit algorithm is upper bounded as follows:*

$$\mathcal{R}(\overline{T}) \leq \mathcal{O}\left(\frac{\alpha_0\nu\omega f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\phi^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log\left(\overline{T}|\mathcal{V}|\right)} + \sqrt[4]{|\mathcal{V}|\log\left(\overline{T}|\mathcal{V}|\right)}\right) + \frac{1}{A}\log(d|\mathcal{V}|) + \sqrt{|\mathcal{V}|}\right),$$

with

$$A = \frac{3\gamma^2\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}^2_{\mathcal{G}}(\mathcal{C}))}{6\frac{\log(|\mathcal{V}|)}{\sqrt{|\mathcal{V}|}} + \sqrt{2}\gamma}, \quad \gamma = \frac{1}{2}\left(1 + \frac{a_2}{a_1}\right)^{-2}.$$

Our regret is mainly formed of two parts. The first one is the sublinear time-dependent term and represents the bulk of horizon dependence. Interestingly, it decreases as the topological centrality index grows with the graph size, which proves the importance of intra-cluster high connectivity.

The second significant term comes from ensuring the RE condition for the empirical multi-task Gram matrix, and can be interpreted as the number of time steps necessary for it to hold, as pointed out by Oh et al. [2021]. It has a logarithmic dependence in the graph size and in the dimension, which is a characteristic of regret bound of the "lasso type". Also noteworthy is that the regret grows explicitly with $\log(d)$ only in the time-independent term, making our policy useful in high-dimensional settings. Though from Proposition 1 we can expect an implicit dependency on the dimension in the RE constant ϕ. Specifically, the lower bound on ϕ is an open problem that appears unsolved in other lasso based works such as Oh et al. [2021], Cella et al. [2023].

Both the regret bound and the oracle inequality presented in Theorem 1 hold only for the set of graphs that at least satisfy the condition $\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} > 2w(\partial\mathcal{P})$ and even though our results hold for a large set of graphs, the individual role of graph-related constants, encapsulated in $f(\mathcal{G}, \mathbf{\Theta}, \alpha_0)$, is not obvious. By further restricting the set of graphs, we are able to provide an alternate bound

Corollary 1. *Assume* $\frac{w(\partial\mathcal{P})(\psi+2\kappa)}{\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}} \leq \Omega$*, with some positive constant* $\Omega < 1$*, then under assumptions 1 to 4, the expected regret of the Network Lasso Bandit algorithm is upper bounded as follows:*

$$\mathcal{R}(\bar{T}) = \mathcal{O}\Bigg(\frac{1}{\phi^2(1-\Omega)} \frac{w(\partial\mathcal{P}) \max_{\mathcal{C}\in\mathcal{P}} \iota_{\mathcal{G}}(\mathcal{C})}{\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}} \sqrt{\bar{T}} \left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\bar{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}| \log(\bar{T}|\mathcal{V}|)}\right) + \frac{w(\partial\mathcal{P})^2 \max_{\mathcal{C}\in\mathcal{P}} \iota_{\mathcal{G}}(\mathcal{C})}{(1-\Omega)^2 \min_{\mathcal{C}\in\mathcal{P}} (\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}^2_{\mathcal{G}}(\mathcal{C}))} \log(d|\mathcal{V}|) + \sqrt{|\mathcal{V}|}\Bigg)$$

The simplified bound in Corollary 1 exhibits the typical multi-task learning dependency $\sqrt{\bar{T}|\mathcal{V}|}$ rather than the independent task learning case $|\mathcal{V}|\sqrt{\bar{T}}$ and highlights the role of graph related properties such as the total weight of the boundary, the maximal inner isoperimetric ratio and the minimal topological centrality index. Furthermore with Ω we can see the influence on the regret bound, when $w(\partial\mathcal{P})$ changes relative to $\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}$.

6 Experiments

We compare our algorithm with $\alpha_0 = 1$ to several baselines of the literature. On the one hand, we consider baselines relying on a given graph, GOBLin [Cesa-Bianchi et al., 2013] and GraphUCB [Yang et al., 2020] that use the Laplacian to smooth the preference vectors. On the other hand, we compare to clustering of bandits baselines, namely CLUB [Gentile et al., 2014], SCLUB [Li et al., 2019], OLS-ITL [Bastani et al., 2021] and LOCB [Ban and He, 2021]. We provided CLUB with graph $\mathcal{G}$ rather than a fully connected graph for a fair comparison. We also include the trace norm bandit algorithm [Cella et al., 2023], which is relevant when the number of clusters is smaller than d (we explain this point in the appendix). As a sanity check, we compare to the independent task learning case with LinUCB (LinUcbITL) where each task is solved independently. The graph used is weightless and generated using a stochastic block model to ensure a cluster structure, where an edge is constructed with probability p within clusters and q between clusters. We present the experimental results in Appendix C.1.

7 Conclusion

In this work, we proposed a multi-task bandit framework that solves the case where the task preference vectors are piecewise constant over a graph. To this end, we used the Network Lasso policy to estimate the task parameters, which bypasses explicit clustering procedures. We established a sublinear regret bound and proved a novel oracle inequality that relies on the small size of the boundary and the high value of the topological centrality index of each node within its cluster. Our experimental evaluations highlight the advantage of our method, especially when either the number of dimensions or nodes increases.

References

Yasin Abbasi-Yadkori, Dávid Pál, and Csaba Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, 2011.

Kaito Ariu, Kenshi Abe, and Alexandre Proutiere. Thresholded Lasso Bandit. In *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Yikun Ban and Jingrui He. Local clustering in contextual multi-armed bandits. In *Proceedings of the Web Conference 2021*, 2021.

Hamsa Bastani and Mohsen Bayati. Online Decision Making with High-Dimensional Covariates. *Operations Research*, 2019.

Hamsa Bastani, Mohsen Bayati, and Khashayar Khosravi. Mostly exploration-free algorithms for contextual bandits. *Manage. Sci.*, 2021.

Javier Borge-Holthoefer, Alejandro Rivero, Iñigo García, Elisa Cauhé, Alfredo Ferrer, Darío Ferrer, David Francos, David Iniguez, María Pilar Pérez, Gonzalo Ruiz, et al. Structural and dynamical patterns on online social networks: the spanish may 15th movement as a case study. *PloS one*, 2011.

Peter Bühlmann and Sara van de Geer. *Statistics for high-dimensional data*. Springer Series in Statistics. Springer, Heidelberg, 2011.

Leonardo Cella and Massimiliano Pontil. Multi-task and meta-learning with sparse linear bandits. In *Uncertainty in Artificial Intelligence*, 2021.

Leonardo Cella, Karim Lounici, Grégoire Pacreau, and Massimiliano Pontil. Multi-task representation learning with stochastic linear bandits. In *International Conference on Artificial Intelligence and Statistics*, 2023.

Nicolo Cesa-Bianchi, Claudio Gentile, and Giovanni Zappella. A gang of bandits. *Advances in neural information processing systems*, 2013.

Jeff Cheeger. A lower bound for the smallest eigenvalue of the laplacian. *Problems in analysis*, 1970.

Xiaotong Cheng, Cheng Pan, and Setareh Maghsudi. Parallel online clustering of bandits via hedonic game. In *International Conference on Machine Learning*, 2023.

Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 2011.

Xiaowen Dong, Dorina Thanou, Michael Rabbat, and Pascal Frossard. Learning graphs from data: A signal representation perspective. *IEEE Signal Processing Magazine*, 2019.

David Easley, Jon Kleinberg, et al. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge university press Cambridge, 2010.

Angela Fontan and Claudio Altafini. On the properties of laplacian pseudoinverses. In *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021.

Claudio Gentile, Shuai Li, and Giovanni Zappella. Online clustering of bandits. In *International Conference on Machine Learning*, 2014.

Claudio Gentile, Shuai Li, Purushottam Kar, Alexandros Karatzoglou, Giovanni Zappella, and Evans Etrue. On context-dependent clustering of bandits. In *International Conference on machine learning*, 2017.

David Hallac, Jure Leskovec, and Stephen Boyd. Network lasso: Clustering and optimization in large graphs. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, 2015.

Xiao He, Francesco Alesiani, and Ammar Shaker. Efficient and scalable multi-task regression on massive number of tasks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019.

Mark Herbster, Stephen Pasteris, Fabio Vitale, and Massimiliano Pontil. A gang of adversarial bandits. *Advances in Neural Information Processing Systems*, 2021.

Daniel Hsu, Sham Kakade, and Tong Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 2012.

Jiachen Hu, Xiaoyu Chen, Chi Jin, Lihong Li, and Liwei Wang. Near-optimal representation learning for linear bandits and linear rl. In *International Conference on Machine Learning*, 2021.

Alexander Jung. Networked Exponential Families for Big Data Over Networks. *IEEE Access*, 2020.

Alexander Jung and Natalia Vesselinova. Analysis of network lasso for semi-supervised regression. In *The 22nd International Conference on Artificial Intelligence and Statistics*, 2019.

Alexander Jung, Nguyen Tran, and Alexandru Mara. When Is Network Lasso Accurate? *Frontiers in Applied Mathematics and Statistics*, 2018.

Gi-Soo Kim and Myunghee Cho Paik. Doubly-robust lasso bandit. *Advances in Neural Information Processing Systems*, 2019.

Lihong Li, Wei Chu, John Langford, and Robert E Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th international conference on World wide web*, 2010.

Shuai Li, Wei Chen, and Kwong-Sak Leung. Improved algorithm on online clustering of bandits. *arXiv preprint arXiv:1902.09162*, 2019.

Miller McPherson, Lynn Smith-Lovin, and James M Cook. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 2001.

Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 2006.

Trong T Nguyen and Hady W Lauw. Dynamic clustering of contextual multi-armed bandits. In *Proceedings of the 23rd ACM international conference on conference on information and knowledge management*, 2014.

Behzad Nourani-Koliji, Steven Bilaj, Amir Rezaei Balef, and Setareh Maghsudi. Piecewise-stationary combinatorial semi-bandit with causally related rewards. In *ECAI 2023*. IOS Press, 2023.

Min-Hwan Oh, Garud Iyengar, and Assaf Zeevi. Sparsity-Agnostic Lasso Bandit. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

Gyan Ranjan and Zhi-Li Zhang. Geometry of complex networks and topological centrality. *Physica A: Statistical Mechanics and its Applications*, 2013.

Xiaoyuan Su and Taghi M Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996.

Joel Tropp. Freedman's inequality for matrix martingales. *Electronic Communications in Probability*, 2011.

Piet Van Mieghem, Karel Devriendt, and H Cetinay. Pseudoinverse of the laplacian and best spreader node in a network. *Physical Review E*, 2017.

Yu-Xiang Wang, James Sharpnack, Alexander J. Smola, and Ryan J. Tibshirani. Trend filtering on graphs. *Journal of Machine Learning Research*, 2016.

Kaige Yang and Laura Toni. Graph-based recommendation system. In *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018.

Kaige Yang, Laura Toni, and Xiaowen Dong. Laplacian-regularized graph bandits: Algorithms and theoretical analysis. In *International Conference on Artificial Intelligence and Statistics*, 2020.

Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 2006.

A Some helper results

Proposition 2 (Bounds on norms of matrix products)**.** *Let $\boldsymbol{M} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{N} \in \mathbb{R}^{n\times p}$. Then*

$$\begin{aligned}\|\boldsymbol{MN}\|_{q,1} &\le \|\boldsymbol{M}\|_{\infty,1}\|\boldsymbol{N}\|_{q,1} \quad \forall q \in [1,\infty]\\ \|\boldsymbol{MN}\|_F &\le \|\boldsymbol{M}\|\|\boldsymbol{N}\|_F\\ \|\boldsymbol{MN}\|_F &\le \sqrt{\|\boldsymbol{M}^\top\boldsymbol{M}\|_{\infty,\infty}}\|\boldsymbol{N}\|_{2,1}\\ \|\boldsymbol{MN}\|_{2,1} &\le \|\boldsymbol{M}\|_{2,1}\|\boldsymbol{N}\|\end{aligned}$$

Proof.

First inequality For any $q \in [1,\infty]$, we have:

$$\left\|\boldsymbol{e}_i^\top\boldsymbol{MN}\right\|_q = \left\|\boldsymbol{e}_i^\top\boldsymbol{M}\sum_{j=1}^n \boldsymbol{e}_j\boldsymbol{e}_j^\top\boldsymbol{N}\right\|_q \le \max_{1\le j\le n}\left|\boldsymbol{e}_i^\top\boldsymbol{M}\boldsymbol{e}_j\right|\sum_{j=1}^n\left\|\boldsymbol{e}_j^\top\boldsymbol{N}\right\|_q = \max_{1\le j\le n}|(\boldsymbol{M})_{ij}|\|\boldsymbol{N}\|_{q,1}$$

Second inequality We have

$$\|\boldsymbol{MN}\|_F^2 = \sum_{j=1}^p\|\boldsymbol{MN}\boldsymbol{e}_j\|^2 \le \sum_{j=1}^p\|\boldsymbol{M}\|\|\boldsymbol{N}\boldsymbol{e}_j\|^2 = \|\boldsymbol{M}\|\|\boldsymbol{N}\|_F^2$$

Third inequality We have

$$\|\boldsymbol{MN}\|_F^2 = \mathrm{Tr}(\boldsymbol{MNN}^\top\boldsymbol{M}^\top) \le \left\|\boldsymbol{M}^\top\boldsymbol{M}\right\|_{\infty,\infty}\left\|\boldsymbol{NN}^\top\right\|_{1,1}$$

Elements of (i,j) entry of matrix $\boldsymbol{NN}^\top$ is the inner product $\left\langle\boldsymbol{e}_i^\top\boldsymbol{N}, \boldsymbol{e}_j^\top\boldsymbol{N}\right\rangle$. Hence, we have

$$\left\|\boldsymbol{NN}^\top\right\|_{1,1} = \sum_{i,j}\left|\left\langle\boldsymbol{e}_i^\top\boldsymbol{N}, \boldsymbol{e}_j^\top\boldsymbol{N}\right\rangle\right| \le \sum_{i,j}\left\|\boldsymbol{e}_i^\top\boldsymbol{N}\right\|\left\|\boldsymbol{e}_j^\top\boldsymbol{N}\right\| = \|\boldsymbol{N}\|_{2,1}^2.$$

Fourth inequality We have

$$\|\boldsymbol{MN}\|_{2,1} = \sum_{i=1}^m\|\boldsymbol{e}_i\boldsymbol{MN}\| \le \sum_{i=1}^m\|\boldsymbol{e}_i\boldsymbol{M}\|\|\boldsymbol{N}\| = \|\boldsymbol{M}\|_{2,1}\|\boldsymbol{N}\|$$

□

Proposition 3 (Decomposition of a signal over a graph)**.** *For any $\mathcal{C} \in \mathcal{P}$*

- *Let $\boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}|\times d}$ be a graph signal. Let us denote by $\boldsymbol{Z}_\mathcal{C}$ the signal obtained from $\boldsymbol{Z}$ by setting rows of vertices outside of $\mathcal{C}$ to zeros, and let $\boldsymbol{Z}_{|\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|\times d}$ be the signal obtained from $\boldsymbol{Z}_\mathcal{C}$ by removing the rows of vertices outside of $\mathcal{C}$. Also, let $\boldsymbol{B}_{|\mathcal{C}} \in \mathbb{R}^{|\mathcal{E}_\mathcal{C}|\times|\mathcal{C}|}$ be the matrix obtained by taking $\boldsymbol{B}_\mathcal{C}$, and removing rows of edges that link $\mathcal{C}$ to its outside, and the resulting null columns. It is clear that*

$$\boldsymbol{B}_\mathcal{C}\boldsymbol{Z} = \boldsymbol{B}_\mathcal{C}\boldsymbol{Z}_\mathcal{C} = \boldsymbol{B}_{|\mathcal{C}}\boldsymbol{Z}_{|\mathcal{C}} \tag{10}$$

- *Let $\boldsymbol{Q}_\mathcal{C} := \boldsymbol{B}_\mathcal{C}^\dagger\boldsymbol{B}_\mathcal{C}$. Then*

$$\boldsymbol{I}_{|\mathcal{V}|} = \sum_{\mathcal{C}\in\mathcal{P}}\boldsymbol{J}_\mathcal{C} + \boldsymbol{Q}_\mathcal{C} \tag{11}$$

$$\boldsymbol{Q}_{\partial\mathcal{P}^c} := \boldsymbol{B}_{\partial\mathcal{P}^c}^\dagger\boldsymbol{B}_{\partial\mathcal{P}^c} = \sum_{\mathcal{C}\in\mathcal{P}}\boldsymbol{Q}_\mathcal{C} \tag{12}$$

where $\boldsymbol{J}_{\mathcal{C}} = \frac{\mathbf{1}_{\mathcal{C}}\mathbf{1}_{\mathcal{C}}^{\top}}{|\mathcal{C}|}$, $\boldsymbol{Q}_{\mathcal{C}} = \boldsymbol{B}_{\mathcal{C}}^{\dagger}\boldsymbol{B}_{\mathcal{C}} \qquad \forall \mathcal{C} \in \mathcal{P}$ *and* $\boldsymbol{Q}_{\partial\mathcal{P}^c} := \boldsymbol{B}_{\partial\mathcal{P}^c}^{\dagger}\boldsymbol{B}_{\partial\mathcal{P}^c}$.

While $\sum_{\mathcal{C}\in\mathcal{P}} \boldsymbol{J}_{\mathcal{C}}$ *projects each entry of a graph signal onto the mean vector value of its respective cluster, its residual* $\boldsymbol{Q}_{\partial\mathcal{P}^c}$ *can be interpreted as the projection onto the respective entries deviation from its cluster mean value.*

Proof. Since the proof of the first point is trivial, we directly treat the second point. Denoting $\boldsymbol{B}_{|\mathcal{C}}^{\dagger}$ the pseudo-inverse of $\boldsymbol{B}_{|\mathcal{C}}$ it is a well-known linear algebra result that the matrix $Q_{|\mathcal{C}} := \boldsymbol{B}_{|\mathcal{C}}^{\dagger}\boldsymbol{B}_{|\mathcal{C}}$ is the projector onto the null space of $\boldsymbol{B}_{|\mathcal{C}}$. Since $\mathcal{C}$ is connected, the null space of $\boldsymbol{B}_{|\mathcal{C}}$ is unidimensional, and is generated by vector $\mathbf{1}_{|\mathcal{C}|} \in \mathbb{R}^{|\mathcal{C}|}$ having only ones as coordinates. Since the projector into that null space is $\boldsymbol{J}_{|\mathcal{C}|} := \frac{\mathbf{1}_{|\mathcal{C}|}\mathbf{1}_{|\mathcal{C}|}}{|\mathcal{C}|}$, we deduce that

$$\begin{aligned} \boldsymbol{Z}_{|\mathcal{C}} &= \boldsymbol{J}_{|\mathcal{C}|}\boldsymbol{Z}_{|\mathcal{C}} + \boldsymbol{Q}_{|\mathcal{C}}\boldsymbol{Z}_{|\mathcal{C}} \\ \Longrightarrow \boldsymbol{Z}_{\mathcal{C}} &= \boldsymbol{J}_{\mathcal{C}}\boldsymbol{Z}_{\mathcal{C}} + \boldsymbol{Q}_{\mathcal{C}}\boldsymbol{Z}_{\mathcal{C}} \\ &= \boldsymbol{J}_{\mathcal{C}}\boldsymbol{Z} + \boldsymbol{Q}_{\mathcal{C}}\boldsymbol{Z} \end{aligned}$$

where in the last line, $Q_{\mathcal{C}} := \boldsymbol{B}_{\mathcal{C}}^{\dagger}\boldsymbol{B}_{\mathcal{C}}$. Consequently, we have

$$\begin{aligned} \boldsymbol{Z} &= \sum_{\mathcal{C}\in\mathcal{P}} \boldsymbol{Z}_{\mathcal{C}} \\ &= \sum_{\mathcal{C}\in\mathcal{P}} \boldsymbol{J}_{\mathcal{C}}\boldsymbol{Z} + \boldsymbol{Q}_{\mathcal{C}}\boldsymbol{Z} \end{aligned}$$

To prove the second point, we recall that $\boldsymbol{B}_{\partial\mathcal{P}^c}$ is the incidence matrix obtained by setting rows corresponding to edges in $\partial\mathcal{P}$ to zero. In other words, $\boldsymbol{B}_{\partial\mathcal{P}^c}$ is the incidence matrix of the graph after removing the boundary edges, and having exactly $|\mathcal{P}|$ connected components. Hence, $\boldsymbol{B}_{\partial\mathcal{P}^c}$ has a null space spanned by the set $\{\mathbf{1}_{\mathcal{C}}\}_{\mathcal{C}\in\mathcal{P}}$, and the orthogonal projector onto this null space is $\sum_{\mathcal{C}\in\mathcal{P}} \boldsymbol{J}_{\mathcal{C}}$. Combining this fact with the fact that $\boldsymbol{Q}_{\partial\mathcal{P}^c}$ is the projector onto the orthogonal of the null space of $\boldsymbol{B}_{\partial\mathcal{P}^c}$, we arrive at the second point. ☐

Proposition 4 (On the minimum topological centrality index of a graph vertex)**.** *Let* $\mathcal{G}$ *be a connected graph with incidence matrix* $\boldsymbol{B}$ *and vertex set size* N*, and let* $\boldsymbol{L} := \boldsymbol{B}^{\top}\boldsymbol{B}$*. Let* $c(\mathcal{G})$ *denote the minimum value of inverses of diagonal element of* $\boldsymbol{L}^{\dagger}$*, called its minimum topological centrality index. Also let* $a(\mathcal{G})$ *be its algebraic connectivity, defined as the minimum non null eigenvalue of* $\boldsymbol{L}$*. Then*

- $c(\mathcal{G}) = \|\boldsymbol{L}\|_{\infty,\infty}^{-1}$.
- $c(\mathcal{G}) \geq a(\mathcal{G})$.
- *If* $\mathcal{G}$ *is weightless, then* $c(\mathcal{G}) \leq \frac{N^2}{N-1}$.

Proof. Since $\boldsymbol{L}$ is PSD, $\boldsymbol{L}^{\dagger}$ is PSD and hence $\left\|\boldsymbol{L}^{\dagger}\right\|_{\infty,\infty}$ is equal to the maximum diagonal entry of $\boldsymbol{L}^{\dagger}$. Taking the inverse proves the first point. Also, this implies that

$$c(\mathcal{G}) = \left\|\boldsymbol{L}^{\dagger}\right\|_{\infty,\infty}^{-1} \geq \left\|\boldsymbol{L}^{\dagger}\right\|^{-1} = a(\mathcal{G}), \tag{13}$$

where we used the fact that $\|\cdot\|_{\infty,\infty} \leq \|\cdot\|$ for matrices. This proves the second point of the proposition.

For the last point, assume $\mathcal{G}$ is weightless, let $\boldsymbol{L}_{\text{comp}}$ be the Laplacian of complete graph built on the vertices of $\mathcal{G}$. Then we have $\boldsymbol{L}_{\text{comp}} = N(\boldsymbol{I}_N - \boldsymbol{J}_N)$, where J is the square matrix of dimension N having $1/N$ as entries. From Fontan and Altafini [2021, Lemma 4], we have

$$\boldsymbol{L}_{\text{comp}}^{\dagger} = (\boldsymbol{L}_{\text{comp}} + N\boldsymbol{J}_N)^{-1} - \frac{1}{N}\boldsymbol{J}_N = \frac{\boldsymbol{I}_N}{N} - \frac{1}{N}\boldsymbol{J}_N \tag{14}$$

which has diagonal elements $\frac{1}{N} - \frac{1}{N^2}$.

On the other hand, $\boldsymbol{L} \preccurlyeq \boldsymbol{L}_{\text{comp}}$ Hence, by Fontan and Altafini [2021, lemma 4] we have for any $u \neq 0$

$$\boldsymbol{L}^{\dagger} = (\boldsymbol{L} + a\boldsymbol{J}_N)^{-1} - \boldsymbol{J}_N/a \succcurlyeq (\boldsymbol{L}_{\text{comp}} + a\boldsymbol{J}_N)^{-1} - \boldsymbol{J}_N/a = \boldsymbol{L}_{\text{comp}}^{\dagger}$$

This implies that the maximum diagonal entry of $\boldsymbol{L}^{\dagger}$ is at least equal to that of $\boldsymbol{L}_{\text{comp}}^{\dagger}$, *i.e.*to $\frac{1}{N} - \frac{1}{N^2}$. Taking the inverse of that entry finishes the proof.

□

B Proofs of the different claims

B.1 Additional notation

The regularization term can be written more compactly using the incidence matrix of the graph $\boldsymbol{B} \in \mathbb{R}^{|\mathcal{E}|\times|\mathcal{V}|}$ corresponding to an arbitrary orientation under the following form

$$\sum_{1\leq m<n\leq|\mathcal{V}|} w_{mn}\|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\| = \|\boldsymbol{B}\boldsymbol{\Theta}\|_{2,1} = \|\boldsymbol{\Theta}\|_{\mathcal{E}} \tag{15}$$

where the $\|\cdot\|_{2,1}$ norm denotes the sum of the L_2 norms o the rows of a matrix.[2] We provide notations that we use in the proofs of the different statements, in order to reduce the clutter. We define $\boldsymbol{E} := \hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}$ as the error signal and its rows by $\{\boldsymbol{\epsilon}_m\}_{m=1}^{|\mathcal{V}|}$.

While $\sum_{k=1}^{C} \boldsymbol{J}_{\mathcal{C}}$ projects each entry of a graph signal onto the mean vector value of its respective cluster, its residual $\boldsymbol{Q}_{\partial\mathcal{P}^c}$ can be interpreted as the projection onto the respective entries deviation from its cluster mean value.

Let $\boldsymbol{\eta}_m$ be a vector, vertically concatenated by noise terms of rewards received by node m, then we define $\boldsymbol{K} \in \mathbb{R}^{|\mathcal{V}|\times d}$ as the matrix of vertically concatenated row vectors $\boldsymbol{\eta}_m^{\top}\boldsymbol{X}_m$.

N.B. Except for the results concerning the regret bound, we consider the case $\kappa \geq 0$ rather than $\kappa \geq 1$ in our proofs.

B.2 Oracle inequality

In this section, we present all intermediary theoretical results leading to Theorem 5 stating the oracle inequality. To reduce clutter, we omit the dependence on t of several quantities. For instance, we write α and $\hat{\boldsymbol{\Theta}}$ instead of $\alpha(t)$ and $\hat{\boldsymbol{\Theta}}(t)$.

Definition 3 (Restricted Eigenvalue (RE) condition and norm, generalization of Definition 2)**.** *Let $\{\boldsymbol{M}_i\}_{i=1}^{|\mathcal{V}|} \subset \mathbb{R}^{d\times d}$ be a set of positive semi-definite matrices. We say that the matrix $\boldsymbol{M}_{\mathcal{V}} := \operatorname{diag}(\boldsymbol{M}_1, \cdots, \boldsymbol{M}_{|\mathcal{V}|})$ verifies the restricted eigenvalue condition with constants $\kappa \geq 0$ and $\phi > 0$ if*

$$\phi^2\|\boldsymbol{Z}\|_{\text{RE}}^2 \leq \sum_{i\in\mathcal{V}} \|\boldsymbol{z}_i\|_{\boldsymbol{M}_i}^2 \quad \forall \boldsymbol{Z} \in \mathcal{S} \text{ with rows } \{\boldsymbol{z}_i\}_{i\in\mathcal{V}},$$

where $\mathcal{S}$ is the cone defined by:

$$\mathcal{S} := \{\boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}|\times d}; a_1\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right)\|\boldsymbol{Z}\|_{\partial\mathcal{P}^c} \leq a_2\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right)\left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F + (1-\kappa)^+\|\boldsymbol{Z}\|_{\partial\mathcal{P}}\},$$

$$a_1(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := 1 - \frac{\frac{1}{\alpha_0} + 2\kappa w(\partial\mathcal{P})}{\min\limits_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}}, \quad a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := \frac{1}{\alpha_0} + \sqrt{2}\kappa w(\partial\mathcal{P}) \max_{\mathcal{C}\in\mathcal{P}} \sqrt{\iota_{\mathcal{G}}(\mathcal{C})},$$

and the RE semi-norm is defined by $\|\boldsymbol{Z}\|_{\text{RE}} := \left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F \vee (1-\kappa)^+\left\|\boldsymbol{B}_{\partial\mathcal{P}}^{\dagger}\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{Z}\right\|$ We have the structure dependent unknown constants ψ and κ, for which we assume they guarantee $a_1\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right) > 0$.

[2]It is possible that the notation $\|\cdot\|_{2,1}$ denotes the sum of 2−norms of columns in the literature.

Table 1: Notation table.

Notation	**Meaning**
	Independent of time t
$\mathcal{V}$	set of graph vertices
$\mathcal{E}$	set of graph edges
$\boldsymbol{B}_I \in \mathbb{R}^{\lvert\mathcal{E}\rvert\times\lvert\mathcal{V}\rvert}, I \subseteq \mathcal{E}$	Graph incidence Matrix obtained by setting rows of edges outside I to zeros
$\boldsymbol{B}_{\mathcal{C}} \in \mathbb{R}^{\lvert\mathcal{E}\rvert\times\lvert\mathcal{V}\rvert}$	cf. Definition 1
$\boldsymbol{L} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times\lvert\mathcal{V}\rvert}$	$\boldsymbol{B}^\top \boldsymbol{B}$
$\boldsymbol{\theta}_m \in \mathbb{R}^d$	true preference vector of user/bandit m
$\boldsymbol{\Theta} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times d}$	matrix of true vertically concatenated row preferences vectors
$\partial\mathcal{P} \subseteq \mathcal{E}$	Boundary of $\mathcal{P}$: set of edges connecting nodes from different clusters
$c_{\mathcal{G}}(\mathcal{C})$	Minimum topological centrality index of a node in $\mathcal{C}$ restricted to the graph in $\mathcal{C}$
$w(\partial\mathcal{P})$	Total weight of $\partial\mathcal{P}$, *i.e.* sum of weights of edges in $\mathcal{P}$
$\lVert\cdot\rVert$	Euclidean norm for vectors, largest singular value for matrices
$\lVert\cdot\rVert_{\boldsymbol{A}}$	Semi-norm defined by PSD matrix $\boldsymbol{A}$: $\lVert\boldsymbol{x}\rVert_{\boldsymbol{A}}^2 := \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}$
$\lVert\cdot\rVert_F$	matrix Frobenius norm
$\lVert\cdot\rVert_{p,q}$	q-norm of the vector with coordinates equal to the $p-$norm of rows
$\lVert\cdot\rVert_I, I \subseteq \mathcal{E}$	Total variation norm of signal over edges of I
$\boldsymbol{A}^\dagger$	Moore-Penrose pseudo-inverse of matrix $\boldsymbol{A}$
vec	vectorization operator consisting in concatenating the columns vertically
$\otimes$	Kronecker product
$\mathbf{1}_{\mathcal{C}} \in \mathbb{R}^{\lvert\mathcal{V}\rvert}$	Vector with entries corresponding to vertices in $\mathcal{C}$ equal to 1 and 0 elsewhere
$\boldsymbol{J}_{\mathcal{C}} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times\lvert\mathcal{V}\rvert}$	equal to $\frac{\mathbf{1}_{\mathcal{C}}\mathbf{1}_{\mathcal{C}}^\top}{\lvert\mathcal{C}\rvert}$
$\boldsymbol{Q}_{\mathcal{C}} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times\lvert\mathcal{V}\rvert}$	equal to $\boldsymbol{B}_{\mathcal{C}}^\dagger \boldsymbol{B}_{\mathcal{C}}$
$\boldsymbol{Q}_I \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times\lvert\mathcal{V}\rvert}, I \subseteq \mathcal{E}$	equal to $\boldsymbol{B}_I^\dagger \boldsymbol{B}_I$
$\boldsymbol{e}_k$	elementary vectors of dimension depending on the context
σ	Subgaussianity constant / variance proxy
	Dependent on time t
$\mathcal{T}_m(t)$	set of time steps user m has been encountered before time t
$\hat{\boldsymbol{\theta}}_m \in \mathbb{R}^d$	estimated preference vector of user/bandit m
$\boldsymbol{\epsilon}_m \in \mathbb{R}^d$	estimation error for user/bandit m : $\hat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m$
$\boldsymbol{E} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times d}$	vertical concatenation of row vectors $\boldsymbol{\epsilon}_m$
$\boldsymbol{\eta}_m \in \mathbb{R}^{\lvert\mathcal{T}_m(t)\rvert}$	vector of subgaussian noise of user m
$\boldsymbol{x}(t) \in \mathbb{R}^d$	context vector received at time t
$m(t) \in \mathbb{N}$	user at time t
$\boldsymbol{X}_m \in \mathbb{R}^{\lvert\mathcal{T}_m(t)\rvert\times d}$	data matrix of user m
$\boldsymbol{X} \in \mathbb{R}^{t\times d}$	data matrix of context vectors of all users
$\boldsymbol{A}_m \in \mathbb{R}^{d\times d}$	$\boldsymbol{X}_m^\top \boldsymbol{X}_m$ (potentially associated to time t)
$\boldsymbol{A}_{\mathcal{V}} \in \mathbb{R}^{d\lvert\mathcal{V}\rvert\times d\lvert\mathcal{V}\rvert}$	$\mathrm{diag}(\boldsymbol{A}_1, \cdots, \boldsymbol{A}_m)$
$\boldsymbol{K} \in \mathbb{R}^{\lvert\mathcal{V}\rvert\times d}$	matrix of vertically concatenated row vectors $\boldsymbol{\eta}_m^\top \boldsymbol{X}_m$

Proof of Proposition 1. Let $\boldsymbol{Z} = \mathbf{1}_{\mathcal{C}}\boldsymbol{v}^\top$ be a constant per cluster signal, with $\mathbf{1}_{\mathcal{C}} \in \mathbb{R}^{\lvert\mathcal{V}\rvert}$ as indicator vector with the ith entry equal to 1 if $i \in \mathcal{C}$ and 0 otherwise. Then $\boldsymbol{Z}$ is contained in any cone $\mathcal{S}$ defined in the RE condition and we have:

$$\left\|\overline{\boldsymbol{Z}_{\mathcal{C}}}\right\|_F^2 = \|\boldsymbol{Z}\|_F^2 = \left\|\mathbf{1}_{\mathcal{C}}\boldsymbol{v}^\top\boldsymbol{v}\mathbf{1}_{\mathcal{C}}^\top\right\| = \mathbf{1}_{\mathcal{C}}^\top\mathbf{1}_{\mathcal{C}}\|\boldsymbol{v}\|^2 = |\mathcal{C}|\|\boldsymbol{v}\|^2$$

For the right hand side of the RE condition we have:

$$\begin{aligned}
\sum_{i\in\mathcal{V}}\|\boldsymbol{z}_i\|_{\boldsymbol{M}_i}^2 &= \operatorname{vec}(\boldsymbol{Z}^\top)^\top\boldsymbol{M}\operatorname{vec}(\boldsymbol{Z}^\top) \\
&= \operatorname{vec}(\mathbf{1}_{\mathcal{C}}\boldsymbol{v}^\top)^\top\boldsymbol{M}\operatorname{vec}(\mathbf{1}_{\mathcal{C}}\boldsymbol{v}^\top) \\
&= (\mathbf{1}_{\mathcal{C}}\otimes\boldsymbol{v})^\top\left(\sum_{i\in\mathcal{V}}\boldsymbol{e}_i\boldsymbol{e}_i^\top\otimes\boldsymbol{M}_i\right)(\mathbf{1}_{\mathcal{C}}\otimes\boldsymbol{v}) \\
&= \left(\mathbf{1}_{\mathcal{C}}^\top\otimes\boldsymbol{v}^\top\right)\left(\sum_{i\in\mathcal{V}}\boldsymbol{e}_i\boldsymbol{e}_i^\top\otimes\boldsymbol{M}_i\right)(\mathbf{1}_{\mathcal{C}}\otimes\boldsymbol{v}) \\
&= \sum_{i\in\mathcal{V}}\mathbf{1}_{\mathcal{C}}^\top\boldsymbol{e}_i\boldsymbol{e}_i^\top\mathbf{1}_{\mathcal{C}}\otimes\boldsymbol{v}^\top\boldsymbol{M}_i\boldsymbol{v} \\
&= \sum_{i\in\mathcal{C}}\boldsymbol{v}^\top\boldsymbol{M}_i\boldsymbol{v} = \boldsymbol{v}^\top\left(\sum_{i\in\mathcal{C}}\boldsymbol{M}_i\right)\boldsymbol{v}
\end{aligned}$$

Plugging the results into the RE condition, we get:

$$\begin{aligned}
\implies &\quad \phi^2|\mathcal{C}|\|\boldsymbol{v}\|^2 \le \boldsymbol{v}^\top\left(\sum_{i\in\mathcal{C}}\boldsymbol{M}_i\right)\boldsymbol{v} \\
\implies &\quad \phi^2 \le \frac{\boldsymbol{v}^\top\left(\sum_{i\in\mathcal{C}}\boldsymbol{M}_i\right)\boldsymbol{v}}{\|\boldsymbol{v}\|^2|\mathcal{C}|} \\
\implies &\quad \phi \le \sqrt{\lambda_{\min}\left(\frac{\sum_{i\in\mathcal{C}}\boldsymbol{M}_i}{|\mathcal{C}|}\right)}
\end{aligned}$$

□

Lemma 1 (A first deterministic inequality)**.** *Let t be a time step. We have*

$$\frac{1}{2t\alpha}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2 + \|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \le \frac{1}{t\alpha}\langle\boldsymbol{K},\boldsymbol{E}\rangle + \|\boldsymbol{E}\|_{\partial\mathcal{P}} \tag{16}$$

Proof. By optimality of $\hat{\boldsymbol{\Theta}}$, we have

$$\frac{1}{2t}\sum_{m\in\mathcal{V}}\left\|\boldsymbol{X}_m\hat{\boldsymbol{\theta}}_m - \boldsymbol{y}_m\right\|^2 + \alpha\|\boldsymbol{\Theta}\|_{\mathcal{E}} \le \frac{1}{2t}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\theta}_m - \boldsymbol{y}_m\|^2 + \alpha\|\boldsymbol{\Theta}\|_{\mathcal{E}} \tag{17}$$

where the second line holds by definition of the observed rewards.

On the one hand, given a user index $m\in\mathcal{V}$, and since by definition of the observed rewards we have we have for the least squared terms

$$\begin{aligned}
\left\|\boldsymbol{X}_m\hat{\boldsymbol{\theta}}_m - \boldsymbol{y}_m\right\|^2 &= \left\|\boldsymbol{X}_m\hat{\boldsymbol{\theta}}_m - \boldsymbol{X}_m\boldsymbol{\theta}_m - \boldsymbol{\eta}_m\right\|^2 \\
&= \|\boldsymbol{X}_m\boldsymbol{\epsilon}_m - \boldsymbol{\eta}_m\|^2 \\
&= \|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2 + \|\boldsymbol{X}_m\boldsymbol{\theta}_m - \boldsymbol{y}_m\|^2 - \boldsymbol{\eta}_m^\top\boldsymbol{X}_m\boldsymbol{\epsilon}_m
\end{aligned}$$

where we used the fact that $\boldsymbol{y}_m = \boldsymbol{X}_m \boldsymbol{\theta}_m + \boldsymbol{\eta}_m$, which holds by definition of the observed rewards. Summing over the users, and using the definition of $\boldsymbol{K}$, we have

$$\frac{1}{2t} \sum_{m \in \mathcal{V}} \left\| \boldsymbol{X}_m \hat{\boldsymbol{\theta}}_m - \boldsymbol{y}_m \right\|^2 - \frac{1}{2t} \sum_{m \in \mathcal{V}} \| \boldsymbol{X}_m \boldsymbol{\theta}_m - \boldsymbol{y}_m \|^2 = \frac{1}{2t} \sum_{m \in \mathcal{V}} \| \boldsymbol{X}_m \boldsymbol{\epsilon}_m \|^2 - \frac{1}{t} \langle \boldsymbol{K}, \boldsymbol{E} \rangle \tag{18}$$

On the other hand, we have for the estimated preference vectors

$$\begin{aligned} \|\boldsymbol{\Theta}\|_{\mathcal{E}} &= \sum_{(m,n) \in \mathcal{E}} w_{mn} \left\| \hat{\boldsymbol{\theta}}_m - \hat{\boldsymbol{\theta}}_n \right\| \\ &= \sum_{(m,n) \in \partial \mathcal{P}} w_{mn} \left\| \hat{\boldsymbol{\theta}}_m - \hat{\boldsymbol{\theta}}_n \right\| + \sum_{(m,n) \in \partial \mathcal{P}^c} w_{mn} \left\| \hat{\boldsymbol{\theta}}_m - \hat{\boldsymbol{\theta}}_n \right\| \\ &= \left\| \hat{\boldsymbol{\Theta}} \right\|_{\partial \mathcal{P}} + \left\| \hat{\boldsymbol{\Theta}} \right\|_{\partial \mathcal{P}^c}, \end{aligned}$$

For the true ones, and for any $\mathcal{C} \in \mathcal{P}$, let $\mathcal{E}_{\mathcal{C}}$ denote the edges linking the nodes of set of nodes $\mathcal{C}$. It is clear that $\partial \mathcal{P}^c = \bigcup_{\mathcal{C} \in \mathcal{P}} \mathcal{E}_{\mathcal{C}}$ as a disjoint union, hence

$$\begin{aligned} \|\boldsymbol{\Theta}\|_{\mathcal{E}} &= \sum_{(m,n) \in \mathcal{E}} w_{mn} \|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\| \\ &= \sum_{(m,n) \in \partial \mathcal{P}} w_{mn} \|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\| + \sum_{(m,n) \in \partial \mathcal{P}^c} w_{mn} \|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\| \\ &= \|\boldsymbol{\Theta}\|_{\partial \mathcal{P}} + \sum_{\mathcal{C} \in \mathcal{P}} \sum_{(m,n) \in \mathcal{E}_{\mathcal{C}}} w_{mn} \|\boldsymbol{\theta}_m - \boldsymbol{\theta}_n\| \\ &= \|\boldsymbol{\Theta}\|_{\partial \mathcal{P}} \end{aligned}$$

where the last equality holds due to the cluster assumption.

Hence, we have

$$\begin{aligned} \|\boldsymbol{\Theta}\|_{\mathcal{E}} - \|\boldsymbol{\Theta}\|_{\mathcal{E}} &= \|\boldsymbol{\Theta}\|_{\partial \mathcal{P}} - \left\| \hat{\boldsymbol{\Theta}} \right\|_{\partial \mathcal{P}} - \left\| \hat{\boldsymbol{\Theta}} \right\|_{\partial \mathcal{P}^c} \\ &\leq \|\boldsymbol{E}\|_{\partial \mathcal{P}} - \left\| \hat{\boldsymbol{\Theta}} \right\|_{\partial \mathcal{P}^c}, \end{aligned} \tag{19}$$

where the first inequality holds due to the triangle inequality, and the last one since $\|\boldsymbol{\Theta}\|_{\partial \mathcal{P}^c} = 0$. Combining Equations (17) to (19), we obtain the result of the statement. □

In the proof for the oracle inequality, we utilize projection operators on the graph signal, which we define as follows:

While $\sum_{k=1}^{C} \boldsymbol{J}_{\mathcal{C}}$ projects each entry of a graph signal onto the mean vector value of its respective cluster, its residual $\boldsymbol{Q}_{\partial \mathcal{P}^c}$ can be interpreted as the projection onto the respective entries deviation from its cluster mean value.

Lemma 2 (Bounding the error restricted to the boundary)**.** *The total variation of $\boldsymbol{E}$ restricted to the boundary verifies*

$$\|\boldsymbol{E}\|_{\partial \mathcal{P}} \leq w(\partial \mathcal{P}) \left(\sqrt{2} \max_{\mathcal{C} \in \mathcal{P}} \sqrt{\iota_{\mathcal{G}}(\mathcal{C})} \left\| \overline{\boldsymbol{E}}_{\mathcal{P}} \right\|_F + 2 \frac{\|\boldsymbol{E}\|_{\partial \mathcal{P}^c}}{\min_{\mathcal{C} \in \mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}} \right) \tag{20}$$

Proof. The proof relies on a decomposition of the $\|\boldsymbol{E}\|_{\partial \mathcal{P}}$ term from Proposition 3. We have

$$\begin{aligned} \|\boldsymbol{E}\|_{\partial \mathcal{P}} &= \left\| \sum_{\mathcal{C} \in \mathcal{P}} \boldsymbol{J}_{\mathcal{C}} \boldsymbol{E} + \boldsymbol{Q}_{\mathcal{C}} \boldsymbol{E} \right\|_{\partial \mathcal{P}} \\ &= \left\| \overline{\boldsymbol{E}}_{\mathcal{P}} + \boldsymbol{B}^{\dagger}_{\partial \mathcal{P}^c} \boldsymbol{B}_{\partial \mathcal{P}^c} \boldsymbol{E} \right\|_{\partial \mathcal{P}} \\ &\leq \left\| \overline{\boldsymbol{E}}_{\mathcal{P}} \right\|_{\partial \mathcal{P}} + \left\| \boldsymbol{B}^{\dagger}_{\partial \mathcal{P}^c} \boldsymbol{B}_{\partial \mathcal{P}^c} \boldsymbol{E} \right\|_{\partial \mathcal{P}} \end{aligned} \tag{21}$$

where $\overline{\boldsymbol{E}}_{\mathcal{P}}$ is obtained by setting the error signal on every cluster to its mean.

For the first term on the right-hand side, let us denote by $\boldsymbol{\epsilon}_{\mathcal{C}}$ the value of any row of $\overline{\boldsymbol{E}}_{\mathcal{P}}$ belonging to cluster $\mathcal{C}$, which is equal to the mean of errors $\boldsymbol{E}$ over that cluster. Also, we denote by $(\overline{\boldsymbol{E}}_{\mathcal{P}})_{\partial\mathcal{P}}$ the signal obtained from $\overline{\boldsymbol{E}}_{\mathcal{P}}$ by setting its rows corresponding to nodes that are not adjacent to any edge in the boundary $\partial\mathcal{P}$ to zeros. Also, let $\partial_v\mathcal{C}$ denote the inner boundary of set of nodes $\mathcal{C}$,i.e. nodes of $\mathcal{C}$ that connect it to its complementary. Then it holds that:

$$\begin{aligned}
\left\|\overline{\boldsymbol{E}}_{\mathcal{P}}\right\|_{\partial\mathcal{P}} &= \left\|\boldsymbol{B}_{\partial\mathcal{P}}\overline{\boldsymbol{E}}_{\mathcal{P}}\right\|_{2,1} \\
&= \left\|\boldsymbol{B}_{\partial\mathcal{P}}(\overline{\boldsymbol{E}}_{\mathcal{P}})_{\partial\mathcal{P}}\right\|_{2,1} \\
&\le \left\|\boldsymbol{B}_{\partial\mathcal{P}}\right\|_{2,1}\left\|(\overline{\boldsymbol{E}}_{\mathcal{P}})_{\partial\mathcal{P}}\right\| \quad \text{(by Proposition 2)} \\
&\le \left\|\boldsymbol{B}_{\partial\mathcal{P}}\right\|_{2,1}\left\|(\overline{\boldsymbol{E}}_{\mathcal{P}})_{\partial\mathcal{P}}\right\|_F \\
&= \left\|\boldsymbol{B}_{\partial\mathcal{P}}\right\|_{2,1}\sqrt{\sum_{\mathcal{C}\in\mathcal{P}}|\partial_v\mathcal{C}|\left\|\boldsymbol{\epsilon}_{\mathcal{C}}\right\|^2} \\
&= \left\|\boldsymbol{B}_{\partial\mathcal{P}}\right\|_{2,1}\sqrt{\sum_{\mathcal{C}\in\mathcal{P}}\frac{|\partial_v\mathcal{C}|}{|\mathcal{C}|}|\mathcal{C}|\left\|\boldsymbol{\epsilon}_{\mathcal{C}}\right\|^2} \\
&\le \left\|\boldsymbol{B}_{\partial\mathcal{P}}\right\|_{2,1}\max_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\sqrt{\sum_{\mathcal{C}\in\mathcal{P}}|\mathcal{C}|\left\|\boldsymbol{\epsilon}_{\mathcal{C}}\right\|^2} \\
&= \sqrt{2}w(\partial\mathcal{P})\max_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}\mathcal{C}}\left\|\overline{\boldsymbol{E}}_{\mathcal{P}}\right\|_F
\end{aligned} \tag{22}$$

For the second term, we have

$$\begin{aligned}
\left\|\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c}\boldsymbol{B}_{\partial\mathcal{P}^c}\boldsymbol{E}\right\|_{\partial\mathcal{P}} &= \left\|\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c}\boldsymbol{B}_{\partial\mathcal{P}^c}\boldsymbol{E}\right\|_{2,1} \\
&\le \left\|\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c}\right\|_{\infty,1}\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \\
&\le \left\|\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c}\right\|_F\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \\
&\le \left\|(\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c})^\top\boldsymbol{B}^\top_{\partial\mathcal{P}}\right\|_F\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \\
&\le \left\|\boldsymbol{B}^\top_{\partial\mathcal{P}}\right\|_{2,1}\sqrt{\left\|\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c}(\boldsymbol{B}^{\dagger}_{\partial\mathcal{P}^c})^\top\right\|_{\infty,\infty}}\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \quad \text{(by Proposition 2)} \\
&= \frac{\left\|\boldsymbol{B}^\top_{\partial\mathcal{P}}\right\|_{1,1}}{\min\limits_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}. \\
&= 2\frac{w(\partial\mathcal{P})}{\min\limits_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}.
\end{aligned} \tag{23}$$

The result is obtained by combining Equations (21) to (23). □

Theorem 4 (Theorem 2.1 of Hsu et al. [2012]). *At time step t, let $\boldsymbol{A} \in \mathbb{R}^{b\times t}$ where $b \in \mathbb{N}^*$, and let $\boldsymbol{v} \in \mathbb{R}^t$ be a random vector such that for some $\sigma \ge 0$, we have*

$$\mathbb{E}\left[\exp(\langle \boldsymbol{u}, \boldsymbol{v}\rangle)\right] \le \exp\left(\|\boldsymbol{u}\|^2\frac{\sigma^2}{2}\right) \quad \forall \boldsymbol{u} \in \mathbb{R}^t.$$

Then for any $\delta \in (0,1)$, we have with a probability at least $1-\delta$:

$$\|\boldsymbol{A}\boldsymbol{v}\|^2 \le \sigma^2\left(\|\boldsymbol{A}\|_F^2 + 2\left\|\boldsymbol{A}^\top\boldsymbol{A}\right\|_F\sqrt{\log\frac{1}{\delta}} + 2\|\boldsymbol{A}\|^2\log\frac{1}{\delta}\right).$$

Lemma 3 (Empirical process bound)**.** *Let $\boldsymbol{X}_m \in \mathbb{R}^{|\mathcal{T}_m|\times d}$ denotes the matrix of collected context vectors for task $m \in \mathcal{V}$, then, given collected context matrices $\{\boldsymbol{X}_m\}_{m\in\mathcal{V}}$, for any $\delta \in (0,1)$ we have with probability of at least $1-\delta$:*

$$\|\boldsymbol{K}\|_F \leq \frac{\alpha_\delta(t)}{\alpha_0} t,$$

where

$$\alpha_\delta(t) := \frac{\alpha_0 \sigma}{t}\sqrt{t + 2\sqrt{\sum_{m\in\mathcal{V}} |\mathcal{T}_m(t)|^2 \log\frac{1}{\delta}} + 2\max_{m\in\mathcal{V}} |\mathcal{T}_m(t)| \log\frac{1}{\delta}}, \tag{24}$$

Proof. We recall that $\boldsymbol{K} \in \mathbb{R}^{t\times d}$ is the matrix obtained by stacking the row vectors $\boldsymbol{\eta}_m^\top \boldsymbol{X}_m$ vertically. On the one hand, we have

$$\|\boldsymbol{K}\|_F^2 = \sum_{m\in\mathcal{V}} \left\|\boldsymbol{X}_m^\top \boldsymbol{\eta}_m\right\|^2 = \left\|\boldsymbol{X}_\mathcal{V}^\top \boldsymbol{\eta}\right\|^2, \tag{25}$$

where $\boldsymbol{X}_\mathcal{V} := \operatorname{diag}(\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{|\mathcal{V}|}) \in \mathbb{R}^{t\times d|\mathcal{V}|}$.

On the other hand, for any $\boldsymbol{u} = (u_1, \cdots, u_t) \in \mathbb{R}^t$, denoting $P(t) := \exp\left(\sum_{\tau=1}^t u_\tau \eta_\tau\right)$, we have

$$\begin{aligned}
\mathbb{E}\left[P(t)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\{u_t \eta_t\} P(t-1) | \mathcal{F}_{t-1}\right]\right] \quad \text{(by the law of total expectation)}\\
&= \mathbb{E}\left[P(t-1)\mathbb{E}\left[\exp(u_t\eta_t)|\mathcal{F}_{t-1}\right]\right] \quad \text{(because } \{\eta_s\}_{s=1}^{t-1} \text{ are } \mathcal{F}_{t-1} \text{ measurable.)}\\
&\leq \exp\left(\frac{1}{2}\sigma^2 u_t^2\right)\mathbb{E}\left[P(t-1)\right] \quad \text{(by the conditional subgaussianity assumption)}\\
&\leq \prod_{s=1}^t \exp\left(\frac{1}{2}\sigma^2 u_s^2\right) \quad \text{(by induction)}\\
&= \exp\left(\frac{1}{2}\sigma^2 \|\boldsymbol{u}\|^2\right).
\end{aligned} \tag{26}$$

From Equations (25) and (26), we can apply Theorem 4 to matrix $\boldsymbol{X}_\mathcal{V}$ and random vector $\boldsymbol{\eta}$, which implies that with a probability at least $1-\delta$, we have

$$\|\boldsymbol{X}_\mathcal{V}\boldsymbol{\eta}\| \leq \sigma\sqrt{\operatorname{Tr}\left(\sum_{m\in\mathcal{V}} \boldsymbol{A}_m\right) + 2\sqrt{\sum_{m\in\mathcal{V}} \|\boldsymbol{A}_m\|_F^2 \log\frac{1}{\delta}} + 2\max_{m\in\mathcal{V}} \|\boldsymbol{A}_m\| \log\frac{1}{\delta}},$$

where we used the equalities $\|\boldsymbol{X}_\mathcal{V}\|_F = \sum_{m\in\mathcal{V}} \operatorname{Tr}(\boldsymbol{A}_m)$, $\|\boldsymbol{X}_\mathcal{V}\|^2 = \max_{m\in\mathcal{V}} \|\boldsymbol{A}_m\|$ and $\left\|\boldsymbol{X}_\mathcal{V}\boldsymbol{X}_\mathcal{V}^\top\right\|_F^2 = \left\|\boldsymbol{X}_\mathcal{V}^\top\boldsymbol{X}_\mathcal{V}\right\|_F^2 = \sum_{m\in\mathcal{V}} \|\boldsymbol{A}_m\|_F^2$. To arrive the the statement of the theorem, we use the fact that the context vectors have Euclidean norms of at most 1.

□

Proposition 5 (Probabilistic inequality)**.** *With a probability at least $1-\delta$, we have*

$$\frac{1}{2t\alpha}\sum_{m\in\mathcal{V}} \|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2 + a_1(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \leq a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\left\|\overline{\boldsymbol{E}}_\mathcal{P}\right\|_F + (1-\kappa)\|\boldsymbol{E}\|_{\partial\mathcal{P}}, \tag{27}$$

where $0 \leq \kappa < \frac{\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_\mathcal{G}(\mathcal{C})}}{2w(\partial\mathcal{P})}$.

Proof. The proof is a combination of the results of Lemmas 1 to 3. We have

$$\frac{1}{2t\alpha_\delta}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2+\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}$$
$$\le\frac{1}{t\alpha_\delta}\langle\boldsymbol{K},\boldsymbol{E}\rangle+\|\boldsymbol{E}\|_{\partial\mathcal{P}}\quad\text{(by Lemma 1)}$$
$$\le\frac{1}{\alpha_0}\|\boldsymbol{E}\|_F+\kappa\|\boldsymbol{E}\|_{\partial\mathcal{P}}+(1-\kappa)\|\boldsymbol{E}\|_{\partial\mathcal{P}}\quad\text{(by Lemma 3)}$$
$$\le\frac{\|\overline{\boldsymbol{E}}_{\mathcal{P}}\|_F}{\alpha_0}+\frac{\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}}{\alpha_0\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}+\kappa w(\partial\mathcal{P})\left(\sqrt{2}\max_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\|\overline{\boldsymbol{E}}_{\mathcal{P}}\|_F+2\frac{\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}}{\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}\right)+(1-\kappa)\|\boldsymbol{E}\|_{\partial\mathcal{P}},$$

where the last line is an application of Lemma 2. Grouping the terms by the type of norm applied to $\boldsymbol{E}$ finishes the proof. □

Theorem 5 (Oracle inequality, generalization of Theorem 1)**.** *Assume that the RE assumption holds for the empirical multi-task Gram matrix with constants* $\kappa\in\left[1,\frac{1}{2w(\partial\mathcal{P})}\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}\right)$, $\psi\in\left(0,\frac{1}{w(\partial\mathcal{P})}\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}-2\right)$ *and* $\phi>0$*. Suppose that* $\max_{m\in\mathcal{V}}|\mathcal{T}_m(t)|\le bt$ *for some* $b>0$ *and* $\alpha_0\ge\frac{1}{\psi w(\partial\mathcal{P})}$*. Then, with a probability at least* $1-\delta(t)$*, we have*

$$\left\|\boldsymbol{\Theta}-\hat{\boldsymbol{\Theta}}(t)\right\|_F\le2\frac{\sigma\alpha_0}{\phi^2\sqrt{t}}f(\mathcal{G},\boldsymbol{\Theta},\alpha_0)\sqrt{1+2b\sqrt{|\mathcal{V}|\log\frac{1}{\delta(t)}}+2b\log\frac{1}{\delta(t)}},$$

where

$$f(\mathcal{G},\boldsymbol{\Theta}):=\left(a_2(\mathcal{G},\boldsymbol{\Theta})+\sqrt{2}\mathbb{1}_{\le1}(\kappa)w(\partial\mathcal{P})\right)\left(\frac{a_2(\mathcal{G},\boldsymbol{\Theta})+\sqrt{2}\mathbb{1}_{\le1}(\kappa)w(\partial\mathcal{P})}{a_1(\mathcal{G},\boldsymbol{\Theta})\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}+1\right).$$

Proof. Using the previously established results, we obtain

$$\frac{1}{2t}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2+\alpha\|\boldsymbol{E}\|_{\partial\mathcal{P}^c}$$
$$\le\alpha_\delta a_2(\boldsymbol{\Theta},\mathcal{G})\|\boldsymbol{E}_{\mathcal{P}}\|_F+\alpha_\delta(1-\kappa)^+\|\boldsymbol{E}\|_{\partial\mathcal{P}}\quad\text{(by Proposition 5)}$$
$$=\alpha_\delta a_2(\boldsymbol{\Theta},\mathcal{G})\|\boldsymbol{E}_{\mathcal{P}}\|_F+\alpha_\delta(1-\kappa)^+\left\|\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{B}_{\partial\mathcal{P}}^\dagger\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{E}\right\|_{2,1}\quad\text{(by properties of the pseudo-inverse)}$$
$$\le\alpha_\delta a_2(\boldsymbol{\Theta},\mathcal{G})\|\boldsymbol{E}_{\mathcal{P}}\|_F+\alpha_\delta\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1}\mathbb{1}_{\le1}(\kappa)(1-\kappa)^+\left\|\boldsymbol{B}_{\partial\mathcal{P}}^\dagger\boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{E}\right\|\quad\text{(by Proposition 2)}$$
$$\le\alpha_\delta(a_2(\boldsymbol{\Theta},\mathcal{G})+\mathbb{1}_{\le1}(\kappa)\sqrt{2}w(\partial\mathcal{P}))\|\boldsymbol{E}\|_{\mathrm{RE}}\quad\text{(by definition of the }\|\|_{\mathrm{RE}}\text{ norm)}$$
$$\le\alpha\frac{a_2(\boldsymbol{\Theta},\mathcal{G})+\mathbb{1}_{\le1}(\kappa)\sqrt{2}w(\partial\mathcal{P})}{\phi\sqrt{t}}\sqrt{\sum_{m\in\mathcal{V}}\|\boldsymbol{\epsilon}_m\|^2_{\boldsymbol{A}_m}}\quad\text{(using the RE assumption)}$$
$$\le\frac{\beta\alpha_\delta^2(a_2(\boldsymbol{\Theta},\mathcal{G})+\mathbb{1}_{\le1}(\kappa)\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1})^2}{2\phi^2}+\frac{1}{2\beta t}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2,\tag{28}$$

where the last inequality holds for any $\beta>0$, and is a consequence of the property that $uv\le\dfrac{u^2+v^2}{2}$ for any $u,v\in\mathbb{R}$. In the second to last inequality we used the RE assumption, here it is important to mention that the assumption does not hold for any choice of α_0. In the definition of $\mathcal{S}$ i.e. the set matrices for which the RE condition holds, we have $\alpha_0=\frac{1}{\psi w(\partial\mathcal{P})}$. We can also observe that this set is non increasing for the inclusion operator i.e. the RE condition would become weaker, for increasing α_0. Thus for any $\alpha_0\ge\frac{1}{\psi w(\partial\mathcal{P})}$ the respective set of the RE assumption is contained in $\mathcal{S}$

and any matrix contained in the smaller set is automatically contained in $\mathcal{S}$, allowing us to use the RE condition in the proof due to our lower bound on $\alpha_0 \geq \frac{1}{\psi w(\partial\mathcal{P})}$.

As a result, we can bound the norm of $\boldsymbol{Q}_{\partial\mathcal{P}^c}\boldsymbol{E}$ as follows:

$$
\begin{aligned}
\|\boldsymbol{Q}_{\partial\mathcal{P}^c}\boldsymbol{E}\|_F &= \left\|\boldsymbol{B}^\dagger_{\partial\mathcal{P}^c}\boldsymbol{B}_{\partial\mathcal{P}^c}\boldsymbol{E}\right\|_F \\
&\leq \sqrt{\left\|\boldsymbol{L}^\dagger_{\partial\mathcal{P}^c}\right\|_{\infty,\infty}}\|\boldsymbol{E}\|_{\partial\mathcal{P}^c} \\
&\leq \frac{2\alpha_\delta(a_2(\boldsymbol{\Theta},\mathcal{G},\alpha_0) + \mathbb{1}_{\leq 1}(\kappa)\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1})^2}{\phi^2 a_1(\boldsymbol{\Theta},\mathcal{G},\alpha_0)\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}} \quad \text{(Equation (28) with } \beta = 1\text{)}. \qquad (29)
\end{aligned}
$$

We can also bound the norm of $\overline{\boldsymbol{E}}_{\mathcal{P}}$ as follows:

$$
\begin{aligned}
\left\|\overline{\boldsymbol{E}}_{\mathcal{P}}\right\|_F^2 &\leq \frac{1}{t\phi^2}\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{\epsilon}_m\|^2 \quad \text{(by RE assumption on empirical multi-task Gram matrix)} \\
&\leq \frac{4\alpha_\delta^2(a_2(\boldsymbol{\Theta},\mathcal{G},\alpha_0) + \mathbb{1}_{\leq 1}(\kappa)\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1})^2}{\phi^4} \quad \text{(by Equation (28) with } \beta = 2\text{)}. \qquad (30)
\end{aligned}
$$

The result is then obtained by combining Equations (29) and (30) along with using the fact that $\boldsymbol{E} = \overline{\boldsymbol{E}}_{\mathcal{P}} + \boldsymbol{Q}_{\partial\mathcal{P}^c}\boldsymbol{E}$ and the expressions of $a_1(\boldsymbol{\Theta},\mathcal{G},\alpha_0)$ and $a_2(\boldsymbol{\Theta},\mathcal{G},\alpha_0)$, and bounding $\alpha_\delta(t)$ as follows:

$$
\begin{aligned}
\frac{\alpha_\delta(t)^2}{\alpha_0^2} &= \frac{\sigma^2}{t^2}\left(\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\|_F^2 + 2\sqrt{\sum_{m\in\mathcal{V}}\|\boldsymbol{X}_m\boldsymbol{X}_m^\top\|_F^2\log\frac{1}{\delta}} + 2\max_{m\in\mathcal{V}}\|\boldsymbol{X}_m\|^2\log\frac{1}{\delta}\right) \\
&\leq \frac{\sigma^2}{t^2}\left(t + 2\sqrt{\sum_{m\in\mathcal{V}}|\mathcal{T}_m(t)|^2\log\frac{1}{\delta}} + 2\max_{m\in\mathcal{V}}|\mathcal{T}_m(t)|\log\frac{1}{\delta}\right) \\
&\leq \frac{\sigma^2}{t^2}\left(t + 2t\sqrt{\log\frac{1}{\delta}} + 2t\log\frac{1}{\delta}\right) \\
&\leq 2\frac{\sigma^2}{t}\left(1 + \sqrt{\log\frac{1}{\delta}}\right)^2
\end{aligned}
$$

□

B.3 Inheriting the RE condition from the true to the empirical data Gram matrix

B.3.1 From the adapted to the empirical multi-task Gram matrix

Lemma 4 (Bounding a quadratic form using projections)**.** *Let $\boldsymbol{M}_1,\cdots,\boldsymbol{M}_p \in \mathbb{R}^{d\times d}$ be symmetric matrices, and let $\boldsymbol{J} := \frac{1}{p}\boldsymbol{1}\boldsymbol{1}^\top$, and $\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{J}$. Then, for any $\boldsymbol{Z} \in \mathbb{R}^{p\times d}$ with rows $\{\boldsymbol{z}_i\}_{i=1}^p$, we have:*

$$
\left|\sum_{i=1}^p \boldsymbol{z}_i^\top\boldsymbol{M}_i\boldsymbol{z}_i\right| \leq \frac{1}{p}\left\|\sum_{i=1}^p\boldsymbol{M}_i\right\|\|\boldsymbol{Z}\|_{\boldsymbol{J}}^2 + 2\sqrt{\left\|\frac{1}{p}\sum_{i=1}^p\boldsymbol{M}_i^2\right\|}\|\boldsymbol{Z}\|_{\boldsymbol{Q}}\|\boldsymbol{Z}\|_{\boldsymbol{J}} + \max_{1\leq i\leq p}\|\boldsymbol{M}_i\|\|\boldsymbol{Z}\|_{\boldsymbol{Q}}^2
$$

Proof. We have

$$
\begin{aligned}
\left|\sum_{i=1}^p \boldsymbol{z}_i^\top\boldsymbol{M}_i\boldsymbol{z}_i\right| &= \left|\sum_{i=1}^p\bar{\boldsymbol{z}}^\top\boldsymbol{M}_i\bar{\boldsymbol{z}} + 2\sum_{i=1}^p(\boldsymbol{z}_i-\bar{\boldsymbol{z}})^\top\boldsymbol{M}_i\bar{\boldsymbol{z}} + \sum_{i=1}^p(\boldsymbol{z}_i-\bar{\boldsymbol{z}})^\top\boldsymbol{M}_i(\boldsymbol{z}_i-\bar{\boldsymbol{z}})\right| \\
&\leq \left|\bar{\boldsymbol{z}}^\top\sum_{i=1}^p\boldsymbol{M}_i\bar{\boldsymbol{z}}\right| + 2\left|\sum_{i=1}^p\boldsymbol{e}_i^\top\boldsymbol{Q}\boldsymbol{Z}\boldsymbol{M}_i\bar{\boldsymbol{z}}\right| + \left|\sum_{i=1}^p\boldsymbol{e}_i^\top\boldsymbol{Q}\boldsymbol{Z}\boldsymbol{M}_i\boldsymbol{Z}^\top\boldsymbol{Q}\boldsymbol{e}_i\right| \qquad (31)
\end{aligned}
$$

where we used the fact that $\boldsymbol{z}_i - \bar{\boldsymbol{z}} = \boldsymbol{Z}^\top \boldsymbol{e}_i - \boldsymbol{Z}^\top \boldsymbol{J}\boldsymbol{e}_i = \boldsymbol{Z}^\top \boldsymbol{Q}\boldsymbol{e}_i$.

Let us now examine every term on the right-hand side of Equation (31). For the first term, we have

$$\left|\bar{\boldsymbol{z}}^\top \sum_{i=1}^p \boldsymbol{M}_i \bar{\boldsymbol{z}}\right| \le \left\|\sum_{i=1}^p \boldsymbol{M}_i\right\| \|\bar{\boldsymbol{z}}\|^2 = \left\|\frac{1}{p}\sum_{i=1}^p \boldsymbol{M}_i\right\| \|\boldsymbol{Z}\|_{\boldsymbol{J}}^2. \tag{32}$$

For the second term, we have

$$\begin{aligned}
\left|\sum_{i=1}^p \boldsymbol{e}_i^\top \boldsymbol{Q}\boldsymbol{Z}\boldsymbol{M}_i \bar{\boldsymbol{z}}\right| &\le \left\|\sum_{i=1}^p \boldsymbol{M}_i \boldsymbol{Z}^\top \boldsymbol{Q}\boldsymbol{e}_i\right\| \|\bar{\boldsymbol{z}}\| \\
&= \left\|\sum_{i=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i) \operatorname{vec}(\boldsymbol{Z}^\top \boldsymbol{Q})\right\| \|\bar{\boldsymbol{z}}\| \\
&\le \left\|\sum_{i=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i)\right\| \|\operatorname{vec}(\boldsymbol{Z}^\top \boldsymbol{Q})\| \|\bar{\boldsymbol{z}}\| \\
&= \left\|\sum_{i=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i)\right\| \|\boldsymbol{Q}\boldsymbol{Z}\|_F \|\bar{\boldsymbol{z}}\| \\
&= \sqrt{\left\|(\sum_{i=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i))^\top \sum_{i=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i)\right\|} \|\boldsymbol{Q}\boldsymbol{Z}\|_F \|\bar{\boldsymbol{z}}\| \\
&= \sqrt{\left\|\sum_{i=1}^p \sum_{j=1}^p (\boldsymbol{e}_i^\top \otimes \boldsymbol{M}_i))(\boldsymbol{e}_j \otimes \boldsymbol{M}_j)\right\|} \|\boldsymbol{Q}\boldsymbol{Z}\|_F \|\bar{\boldsymbol{z}}\| \\
&= \sqrt{\left\|\sum_{i=1}^p \sum_{j=1}^p (\boldsymbol{e}_i^\top \boldsymbol{e}_j \otimes \boldsymbol{M}_i \boldsymbol{M}_j)\right\|} \|\boldsymbol{Q}\boldsymbol{Z}\|_F \|\bar{\boldsymbol{z}}\| \\
&= \sqrt{\left\|\sum_{i=1}^p \boldsymbol{M}_i^2\right\|} \|\boldsymbol{Q}\boldsymbol{Z}\|_F \|\bar{\boldsymbol{z}}\|.
\end{aligned} \tag{33}$$

Finally, for the last term, we have

$$\begin{aligned}
\left|\sum_{i=1}^p \boldsymbol{e}_i^\top \boldsymbol{Q}\boldsymbol{Z}\boldsymbol{M}_i \boldsymbol{Z}^\top \boldsymbol{Q}\boldsymbol{e}_i\right| &\le \sum_{i=1}^p \|\boldsymbol{M}_i\| \left\|\boldsymbol{Z}^\top \boldsymbol{Q}\boldsymbol{e}_i\right\|^2 \\
&\le \max_{1\le i\le p} \|\boldsymbol{M}_i\| \sum_{i=1}^p \left\|\boldsymbol{Z}^\top \boldsymbol{Q}\boldsymbol{e}_i\right\|^2 \\
&= \max_{1\le i\le p} \|\boldsymbol{M}_i\| \|\boldsymbol{Q}\boldsymbol{Z}\|_F^2.
\end{aligned} \tag{34}$$

Combining Equations (32) to (34) yields the result. □

We also define an operator norm that is induced by the $\|\|_{\mathrm{RE}}$ introduced in Definition 3.

Definition 4 ((RE,$\mathcal{S}$)-induced operator norm)**.** *Let $\{\boldsymbol{M}_m\}_{m\in\mathcal{V}} \subseteq \mathbb{R}^{d\times d}$ be symmetric matrices associated to the graph nodes $\mathcal{V}$, and let $\boldsymbol{M}_{\mathcal{V}} := \operatorname{diag}\left(\boldsymbol{M}_1, \cdots, \boldsymbol{M}_{|\mathcal{V}|}\right) \in \mathbb{R}^{d|\mathcal{V}|\times d|\mathcal{V}|}$. For any cluster $\mathcal{C} \in \mathcal{P}$, let the cluster mean and mean of squares associated to those matrices be given by*

$$\overline{\boldsymbol{M}}_{\mathcal{C}} := \frac{1}{|\mathcal{C}|}\sum_{m\in\mathcal{C}} \boldsymbol{M}_m, \qquad \overline{\boldsymbol{M}^2}_{\mathcal{C}} := \frac{1}{|\mathcal{C}|}\sum_{m\in\mathcal{C}} \boldsymbol{M}_m^2.$$

The RE-induced operator norm of $\boldsymbol{M}_{\mathcal{V}}$ is defined as

$$\|\boldsymbol{M}\|_{\mathrm{RE},\mathcal{S}} := \max_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}}_{\mathcal{C}}\right\| \vee \sqrt{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1} \max_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|} \vee \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1} \max_{m\in\mathcal{V}} \|\boldsymbol{M}_m\|. \tag{35}$$

B.3.2 Linking the adapted to the empirical Gram

We first start by establishing that given the closeness of two PSD matrices in a certain sense, the RE condition can be transferred between them. For the sake of readability we remove the arguments of the constants: $a_1 = a_1\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right)$, $a_2 = a_2\left(\mathcal{G}, \boldsymbol{\Theta}, \frac{1}{\psi w(\partial\mathcal{P})}\right)$,

Proposition 6 (Restricted spectral norm)**.** *Let $\boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}|\times d}$ verifying*

$$a_1\|\boldsymbol{Z}\|_{\partial\mathcal{P}^c} \leq a_2\left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F + (1-\kappa)^+\|\boldsymbol{Z}\|_{\partial\mathcal{P}}$$

Let $\{\boldsymbol{M}_m\}_{m\in\mathcal{V}} \subseteq \mathbb{R}^{d\times d}$ be symmetric matrices associated to the graph nodes $\mathcal{V}$, and let $\boldsymbol{M}_{\mathcal{V}} \coloneqq \operatorname{diag}(\boldsymbol{M}_1, \cdots, \boldsymbol{M}_{|\mathcal{V}|}) \in \mathbb{R}^{d|\mathcal{V}|\times d|\mathcal{V}|}$. Then we have:

$$\left|\sum_{m\in\mathcal{V}} \boldsymbol{z}_m^\top \boldsymbol{M}_m \boldsymbol{z}_m\right| \leq \|\boldsymbol{M}\|_{\mathrm{RE},\mathcal{S}}^2 \left(1 + \frac{a_2 + (1-\kappa)^+\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1}}{a_1}\right)^2 \|\boldsymbol{Z}\|_{\mathrm{RE}}^2. \tag{36}$$

Proof. For any cluster $\mathcal{C}$, we denote by $\boldsymbol{B}_{\mathcal{C}}$ the incidence matrix obtained by setting the rows of $\boldsymbol{B}$ outside the edges linking nodes in $\mathcal{C}$ to null vectors. The latter's null space is the span of the vector $\mathbf{1}_{\mathcal{C}}$ having coordinates 1 at nodes in $\mathcal{C}$ and zeros elsewhere. Hence, the projector onto the orthogonal of $\mathbf{1}_{\mathcal{C}}$ is $\boldsymbol{Q}_{\mathcal{C}} \coloneqq \boldsymbol{B}_{\mathcal{C}}^\dagger \boldsymbol{B}_{\mathcal{C}}$.

On the one hand, for any signal $\boldsymbol{Z} \in \mathbb{R}^{|\mathcal{V}|\times d}$ we have

$$\begin{aligned}
\|\boldsymbol{Z}\|_{\partial\mathcal{P}^c} &= \sum_{\mathcal{C}\in\mathcal{P}} \|\boldsymbol{B}_{\mathcal{C}}\boldsymbol{Z}\|_{2,1} \\
&\geq \sum_{\mathcal{C}\in\mathcal{P}} \frac{\left\|\boldsymbol{B}_{\mathcal{C}}^\dagger \boldsymbol{B}_{\mathcal{C}}\boldsymbol{Z}\right\|_F}{\sqrt{\left\|\boldsymbol{L}_{\mathcal{C}}^\dagger\right\|_{\infty,\infty}}} \\
&\geq \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} \sum_{\mathcal{C}\in\mathcal{P}} \|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}}
\end{aligned}$$

Hence, by the proposition's assumptions, $\boldsymbol{Z}$ verifies

$$\begin{aligned}
\min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}a_1 \sum_{\mathcal{C}\in\mathcal{P}} \|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}} &\leq (a_2\left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F + (1-\kappa)\|\boldsymbol{Z}\|_{\partial\mathcal{P}}) \\
&\leq a_2\left\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\right\|_F + (1-\kappa)^+\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1}\left\|\boldsymbol{B}_{\partial\mathcal{P}}^\dagger \boldsymbol{B}_{\partial\mathcal{P}}\boldsymbol{Z}\right\| \\
&\leq (a_2 + (1-\kappa)^+\|\boldsymbol{B}\|_{2,1})\|\boldsymbol{Z}\|_{\mathrm{RE}}
\end{aligned}$$

From Lemma 4, we have

$$\begin{aligned}
&\left|\sum_{m\in\mathcal{V}} \boldsymbol{z}_m^\top \boldsymbol{M}_m \boldsymbol{z}_m\right| \\
&\leq \sum_{\mathcal{C}\in\mathcal{P}} \left|\sum_{m\in\mathcal{C}} \boldsymbol{z}_m^\top \boldsymbol{M}_m \boldsymbol{z}_m\right| \\
&\leq \sum_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}}_{\mathcal{C}}\right\|\|\boldsymbol{Z}\|_{\boldsymbol{J}_{\mathcal{C}}}^2 + 2\sum_{\mathcal{C}\in\mathcal{P}} \sqrt{\left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|}\|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}}\|\boldsymbol{Z}\|_{\boldsymbol{J}_{\mathcal{C}}} + \sum_{\mathcal{C}\in\mathcal{P}} \max_{m\in\mathcal{C}} \|\boldsymbol{M}_m\|\|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}}^2,
\end{aligned} \tag{37}$$

where we used Equation (10).

This allows us to bound every term in Equation (37). For the second term on the right-hand side, we have

$$\begin{aligned}
&\sum_{\mathcal{C}\in\mathcal{P}} \sqrt{\left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|}\|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}}\|\boldsymbol{Z}\|_{\boldsymbol{J}_{\mathcal{C}}} \\
&\leq \max_{\mathcal{C}\in\mathcal{P}} \sqrt{\left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|}\|\overline{\boldsymbol{Z}}_{\mathcal{P}}\|_F \sqrt{\sum_{\mathcal{C}\in\mathcal{P}} \|\boldsymbol{Z}\|^2_{\boldsymbol{Q}_{\mathcal{C}}}} \\
&\leq \frac{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-\frac{1}{2}}}{a_1} \max_{\mathcal{C}\in\mathcal{P}} \sqrt{\left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|}(a_2 + (1-\kappa)^+\|\boldsymbol{B}\|_{2,1})\|\boldsymbol{Z}\|^2_{\mathrm{RE}} \qquad (38)
\end{aligned}$$

As for the third term, we have

$$\begin{aligned}
\sum_{\mathcal{C}\in\mathcal{P}} \max_{m\in\mathcal{C}} \|\boldsymbol{M}_m\|\|\boldsymbol{Z}\|^2_{\boldsymbol{Q}_{\mathcal{C}}} &\leq \max_{m\in\mathcal{V}} \|\boldsymbol{M}_m\| \left(\sum_{\mathcal{C}\in\mathcal{P}} \|\boldsymbol{Z}\|_{\boldsymbol{Q}_{\mathcal{C}}}\right)^2 \\
&\leq \max_{m\in\mathcal{V}} \|\boldsymbol{M}_m\| \frac{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1}}{a_1^2}(a_2 + (1-\kappa)^+\|\boldsymbol{B}\|_{2,1})^2\|\boldsymbol{Z}\|^2_{\mathrm{RE}} \qquad (39)
\end{aligned}$$

Consequently, denoting $v = \dfrac{a_2 + (1-\kappa)^+\|\boldsymbol{B}\|_{2,1}}{a_1}$, and combining Equations (37) to (39), we obtain

$$\begin{aligned}
&\left|\sum_{m\in\mathcal{V}} \boldsymbol{z}_m^\top \boldsymbol{M}_m \boldsymbol{z}_m\right| \\
&\left(\max_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}}_{\mathcal{C}}\right\| + 2v \max_{\mathcal{C}\in\mathcal{P}} \sqrt{\left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|} + v^2 \max_{i\in\mathcal{V}} \|\boldsymbol{M}_i\|\right) \|\boldsymbol{Z}\|^2_{\mathrm{RE}} \\
&\leq \left(\max_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}}_{\mathcal{C}}\right\|) \vee \sqrt{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1} \max_{\mathcal{C}\in\mathcal{P}} \left\|\overline{\boldsymbol{M}^2}_{\mathcal{C}}\right\|} \vee \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1} \max_{i\in\mathcal{V}} \|\boldsymbol{M}_i\|\right) (1+v)^2\|\boldsymbol{Z}\|^2_{\mathrm{RE}},
\end{aligned}$$

which finishes the proof. □

Proposition 7 (Inheritance of a RE condition from a close matrix)**.** *Assume that the matrix $\boldsymbol{V}_{\mathcal{V}}$ verifies the RE condition with constant $\phi > 0$, and that $\left\|\dfrac{\boldsymbol{A}_{\mathcal{V}}}{t} - \boldsymbol{V}_{\mathcal{V}}\right\|_{\mathrm{op,RE}} \leq \gamma\phi^2$ for some $\gamma \in \left(0, \left(1 + \frac{a_2+(1-\kappa)^+\sqrt{2}w(\partial\mathcal{P})}{a_1}\right)^{-2}\right)$. Then $\dfrac{\boldsymbol{A}_{\mathcal{V}}}{t}$ verifies the RE condition with constant*

$$\hat{\phi} = \phi\sqrt{1 - \gamma\left(1 + \frac{a_2 + (1-\kappa)^+\sqrt{2}w(\partial\mathcal{P})}{a_1}\right)^2} \qquad (40)$$

Proof. From Proposition 5, we know that

$$\begin{aligned}
\frac{1}{t}\epsilon_{\mathcal{V}}^\top \boldsymbol{A}_{\mathcal{V}}\epsilon_{\mathcal{V}} &= \frac{1}{|\mathcal{V}|}\epsilon_{\mathcal{V}}^\top \boldsymbol{V}_{\mathcal{V}}\epsilon_{\mathcal{V}} + \epsilon_{\mathcal{V}}^\top \boldsymbol{\Delta}_{\mathcal{V}}\epsilon_{\mathcal{V}} \\
&\geq \frac{1}{|\mathcal{V}|}\epsilon_{\mathcal{V}}^\top \boldsymbol{V}_{\mathcal{V}}\epsilon_{\mathcal{V}} - \left|\epsilon_{\mathcal{V}}^\top \boldsymbol{\Delta}_{\mathcal{V}}\epsilon_{\mathcal{V}}\right| \\
&\geq \left(\phi^2 - \max_{m\in\mathcal{V}} \|\boldsymbol{\Delta}_{\mathcal{V}}\|_{\mathrm{op,RE}} \left(1 + \frac{a_2 + (1-\kappa)^+\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1}}{a_1}\right)^2\right) \|\boldsymbol{E}\|^2_{\mathrm{RE}} \\
&\geq \left(\phi^2 - \gamma\phi^2 \left(1 + \frac{a_2 + (1-\kappa)^+\|\boldsymbol{B}_{\partial\mathcal{P}}\|_{2,1}}{a_1}\right)^2\right) \|\boldsymbol{E}\|^2_{\mathrm{RE}}
\end{aligned}$$

where the third inequality is an applicaiton of Proposition 6. □

Theorem 6 (Matrix Freedman Inequality, Tropp [2011]). *Consider a matrix martingale $\{\boldsymbol{M}(t)\}_{t\geq 1}$ with dimension $d_1 \times d_2$. Let $\{\boldsymbol{N}(t)\}_{t\geq 1}$ be the associated difference sequence. Assume that for some $A > 0$, we have $\|\boldsymbol{N}(t)\| \leq A \quad \forall t \geq 1$ almost surely. Define for any $t \geq 1$:*

$$\boldsymbol{W}_{col}(t) \coloneqq \sum_{\tau=1}^{t} \mathbb{E}\left[\boldsymbol{N}(\tau)\boldsymbol{N}(\tau)^\top | \mathcal{F}_{\tau-1}\right]$$
$$\boldsymbol{W}_{row}(t) \coloneqq \sum_{\tau=1}^{t} \mathbb{E}\left[\boldsymbol{N}(\tau)^\top \boldsymbol{N}(\tau) | \mathcal{F}_{\tau-1}\right].$$

Then, for any $u, v > 0$,

$$\mathbb{P}\left[\exists t \geq 1; \|\boldsymbol{M}(t)\| \geq u \text{ and } \|\boldsymbol{W}_{col}\|(t) \vee \|\boldsymbol{W}_{row}(t)\| \leq v\right] \leq (d_1 + d_2) \exp\left(-\frac{3u^2}{6v + 2Au}\right)$$

Corollary 2. *Let $\{\boldsymbol{N}(\tau)\}_{\tau=1}^{t}$ by a sequence of matrices of dimension $d_1 \times d_2$, adapted to filtration $\{\mathcal{F}_\tau\}_{\tau=1}^{t}$. Let $\{t_i\}_{i=1}^{N}$ an increasing sequence with elements in $[t]$ for some $N \leq t$. Consider the sequence $\{\boldsymbol{M}(n)\}_{\tau=1}^{N}$ of random matrices defined by*

$$\boldsymbol{M}(n) = \sum_{i=1}^{n} \boldsymbol{N}(t_i) - \mathbb{E}\left[\boldsymbol{N}(t_i) | \mathcal{F}_{t_i - 1}\right] \tag{41}$$

Then $\{\boldsymbol{M}(n)\}_{n=1}^{N}$ is a martingale adapted to the filtration $\{\mathcal{F}_{t_n}\}_{n=1}^{N}$.

Moreover,if $\|\boldsymbol{N}(\tau)\| \leq b \quad \forall \tau \in [t]$ for some $b > 0$, then we have

$$\mathbb{P}\left[\|\boldsymbol{M}(N)\| \geq u\right] \leq (d_1 + d_2) \exp\left(-\frac{3u^2}{6Nb^2 + 2\sqrt{2}bu}\right). \tag{42}$$

Proof. We denote $\mathbb{E}\left[\cdot|\mathcal{F}_s\right]$ as $\mathbb{E}_s\left[\cdot\right]$ for any $s \in \mathbb{N}$. Also, let $\boldsymbol{C}(s) \coloneqq \mathbb{E}_{s-1}\left[\boldsymbol{N}(s)\right]$, which is $\mathcal{F}_{s-1}$-measurable by construction. We have for any $n \in [N]$,

$$\mathbb{E}_{t_{n-1}}\left[\boldsymbol{C}(t_n)\right] = \mathbb{E}_{t_{n-1}}\left[\mathbb{E}_{t_n - 1}\left[\boldsymbol{N}(t_n)\right]\right] = \mathbb{E}_{t_{n-1}}\left[\boldsymbol{N}(t_n)\right] \tag{43}$$
$$\implies \mathbb{E}_{t_{n-1}}\left[\boldsymbol{N}(t_n) - \boldsymbol{C}(t_n)\right] = 0 \tag{44}$$

where the first equality is due to the tower rule since $\mathcal{F}_{t_{n-1}} \subset \mathcal{F}_{t_n - 1}$. Also, we have for any $\tau \geq 1$

$$\|\boldsymbol{N}(\tau) - \boldsymbol{C}(\tau)\|^2 = \left\|(\boldsymbol{N}(\tau) - \boldsymbol{C}(\tau))^2\right\| \tag{45}$$
$$\leq \mathrm{Tr}\big((\boldsymbol{N}(\tau) - \boldsymbol{C}(\tau))^2\big) \tag{46}$$
$$= \mathrm{Tr}\big((\boldsymbol{N}(\tau) - \boldsymbol{C}(\tau))^2\big) \tag{47}$$
$$= \|\boldsymbol{N}(\tau)\|_F^2 - 2\,\mathrm{Tr}(\boldsymbol{C}(\tau)\boldsymbol{N}(\tau)) + \mathrm{Tr}\big(\boldsymbol{C}(\tau)^2\big) \tag{48}$$
$$\leq \|\boldsymbol{N}(\tau)\|_F^2 + \mathrm{Tr}\big(\boldsymbol{C}(\tau)^2\big) \leq 2b^2 \tag{49}$$

Hence $\boldsymbol{N}(\tau) - \boldsymbol{C}(\tau)$ is integrable for any $\tau \geq 1$. This shows that $\boldsymbol{M}(n)$ is a sequence of partial sums of matrix martingale differences, hence it is a matrix martingale.

The second part of the corollary statement is a consequence of Theorem 6. The boundedness of the sequence of martingale differences has already been established above. To verify the second requirement of the theorem, let us compute bounds on the norms of $\boldsymbol{W}_{\text{col}}$ and $\boldsymbol{W}_{\text{row}}$ from Theorem 6. Notice that the two matrices are equal since the difference sequence matrices $\boldsymbol{N}(t_s)$ are symmetric.

Hence, for any $n \in [N]$, we have

$$\|\boldsymbol{W}_{\text{col}}(N)\| \vee \|\boldsymbol{W}_{\text{row}}(N)\| \leq \operatorname{Tr}(\boldsymbol{W}_{\text{col}}(N)) \vee \operatorname{Tr}(\boldsymbol{W}_{\text{row}}(N)) \tag{50}$$

$$= \operatorname{Tr}\left(\sum_{n=1}^{N} \mathbb{E}_{t_n-1}\left[(\boldsymbol{N}(t_n) - \boldsymbol{C}(t_n))^2\right]\right) \tag{51}$$

$$= \sum_{n=1}^{N} \mathbb{E}_{t_n-1}\left[\|\boldsymbol{N}(t_n)\|_F^2\right] - \mathbb{E}_{t_n-1}\left[2\operatorname{Tr}(\boldsymbol{C}(t_n)\boldsymbol{N}(t_n))\right] + \operatorname{Tr}\big(\boldsymbol{C}(t_n)^2\big) \tag{52}$$

$$= \sum_{n=1}^{N} \mathbb{E}_{t_n-1}\left[\|\boldsymbol{N}(t_n)\|_F^2\right] - \operatorname{Tr}\big(\boldsymbol{C}(t_n)^2\big) \tag{53}$$

$$\leq \sum_{n=1}^{N} \mathbb{E}_{t_n-1}\left[\|\boldsymbol{N}(t_n)\|_F^2\right] \leq Nb^2. \tag{54}$$

By Theorem 6, we have for any $u > 0$

$$2d\exp\left(-\frac{3u^2}{6Nb^2 + 2\sqrt{2}bu}\right) \geq \mathbb{P}\left[\exists n \geq 1; \|\boldsymbol{M}(n)\| \geq u \text{ and } \|\boldsymbol{W}_{\text{col}}(n)\| \leq Nb^2\right] \tag{55}$$

$$\geq \mathbb{P}\left[\|\boldsymbol{M}(N)\| \geq u \text{ and } \|\boldsymbol{W}_{\text{col}}(N)\| \leq Nb^2\right] \tag{56}$$

$$= \mathbb{P}\left[\|\boldsymbol{M}(N)\| \geq u\right] \tag{57}$$

where the last line holds because we showed that the inequality $\|\boldsymbol{W}_{\text{col}}(N)\| \leq Nb^2$ holds almost surely. □

Proposition 8 (Concentration of the empirical multi-task Gram matrix around the adapted one)**.** *Let $t \geq 1$, $b > 0$. Then we have:*

$$\mathbb{P}\left[\left\|\frac{\boldsymbol{A}_{\mathcal{V}}(t)}{t} - \boldsymbol{V}_{\mathcal{V}}\right\|_{\text{op,RE}} > \gamma \Big| \max_{m\in\mathcal{V}} |\mathcal{T}_m(t)| \leq bt\right] \leq d(2|\mathcal{P}|e^{-A_1 t} + (|\mathcal{V}| + |\mathcal{P}|)e^{-A_2 t} + 2|\mathcal{V}|e^{-A_3 t}),$$

where

$$A_1 := \frac{3\gamma^2 \min_{\mathcal{C}\in\mathcal{P}} |\mathcal{C}| t}{6b + 2\sqrt{2}\gamma}$$

$$A_2 := \frac{3\gamma^2 \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C}) t}{6b + 2\sqrt{2}\gamma\sqrt{\frac{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}{\min_{\mathcal{C}\in\mathcal{P}} |\mathcal{C}|}}}$$

$$A_3 := \frac{3\gamma^2 \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^2 t}{6b + 2\sqrt{2}\gamma \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}$$

Proof. For $\gamma > 0$, let us define

$$\boldsymbol{\Delta}_m := \frac{\boldsymbol{A}_{\mathcal{V}}}{t} - \boldsymbol{V}_{\mathcal{V}} \quad \text{and } G_{\text{Gram},\gamma} := \left\{\frac{1}{t}\|\boldsymbol{\Delta}_{\mathcal{V}}\|_{\text{RE},\mathcal{S}} \leq \gamma\right\},$$

where $\boldsymbol{\Delta}_{\mathcal{V}}$ is block diagonal matrix formed by $\{\boldsymbol{\Delta}_m\}_{m\in\mathcal{V}}$. We also define $\overline{\boldsymbol{\Delta}}_{\mathcal{C}}$ and $\overline{\boldsymbol{\Delta}^2}_{\mathcal{C}}$ in the same pattern of Definition 4. We can express the complementary of this event as the disjunction of a finite number of events as follows:

$$G^{c}_{\mathrm{Gram},\gamma} \tag{58}$$
$$=\left\{\max_{\mathcal{C}\in\mathcal{P}}\left\|\overline{\boldsymbol{\Delta}}_{\mathcal{C}}\right\| \vee \sqrt{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1}\max_{\mathcal{C}\in\mathcal{P}}\left\|\overline{\boldsymbol{\Delta}^2}_{\mathcal{C}}\right\|} \vee \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^{-1}\max_{m\in\mathcal{V}}\|\boldsymbol{\Delta}_m\| > t\gamma\right\} \tag{59}$$
$$=\bigcup_{\mathcal{C}\in\mathcal{P}}\left\{\left\|\overline{\boldsymbol{\Delta}}_{\mathcal{C}}\right\| > t\gamma\right\} \cup \bigcup_{\mathcal{C}\in\mathcal{P}}\left\{\left\|\overline{\boldsymbol{\Delta}^2}_{\mathcal{C}}\right\| > t^2\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})\right\} \cup \bigcup_{m\in\mathcal{V}}\left\{\|\boldsymbol{\Delta}_m\| > t\gamma\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})\right\} \tag{60}$$

The first and third event can be bounded by considering the sequence $\boldsymbol{x}\boldsymbol{x}^\top(\tau)$ adapted to the filtration $\{\mathcal{F}_\tau\}$, verifying $\left\|\boldsymbol{x}\boldsymbol{x}^\top(\tau)\right\| \leq$.

Bounding the probability of the first event Let $\mathcal{C}\in\mathcal{P}$ be a cluster. By definition, we have

$$\begin{aligned}|\mathcal{C}|\overline{\boldsymbol{\Delta}}_{\mathcal{C}}(t) &= \sum_{m\in\mathcal{C}}\sum_{\tau\in\mathcal{T}_m(t)} \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right] \\ &= \sum_{\tau\in\bigcup_{m\in\mathcal{C}}\mathcal{T}_m(t)} \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]\end{aligned}$$

We will apply Corollary 2 for the sequence of time indices in $\mathcal{C}$, *i.e.* $\bigcup_{m\in\mathcal{V}}\mathcal{T}_m(t)$. Hence $|\mathcal{C}|\overline{\boldsymbol{\Delta}_{\mathcal{C}}}$ is a martingale sequence, and we have

$$\begin{aligned}\mathbb{P}\left[\left\|\overline{\boldsymbol{\Delta}}_{\mathcal{C}}(t)\right\| > \gamma t \,\Big|\, \max_{m\in\mathcal{V}}|\mathcal{T}_m(t)| \leq bt\right] &\leq 2d\exp\left(\frac{-3\gamma^2|\mathcal{C}|^2t^2}{6\sum_{m\in\mathcal{C}}|\mathcal{T}_m(t)| + 2\sqrt{2}\gamma|\mathcal{C}|t}\right) \\ &\leq 2d\exp\left(\frac{-3\gamma^2|\mathcal{C}|^2t^2}{6|\mathcal{C}|bt + 2\sqrt{2}\gamma|\mathcal{C}|t}\right) \\ &= 2d\exp\left(\frac{-3\gamma^2|\mathcal{C}|t}{6b + 2\sqrt{2}\gamma}\right) \\ &\leq 2d\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}}|\mathcal{C}|t}{6b + 2\sqrt{2}\gamma}\right)\end{aligned} \tag{61}$$

Bounding the probability of the third event Let $m\in\mathcal{V}$ be a task index. We apply Corollary 2 for the sequence of time steps in $\mathcal{T}_m(t)$. We have

$$\boldsymbol{\Delta}_m(t) = \sum_{\tau\in\mathcal{T}_m(t)} \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]$$

is a martingale sequence, hence

$$\begin{aligned}\mathbb{P}\left[\|\boldsymbol{\Delta}_m(t)\| > \gamma\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t \,\Big|\, \max_{m\in\mathcal{V}}|\mathcal{T}_m(t)| \leq bt\right] &\leq 2d\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^2t^2}{6|\mathcal{T}_m(t)| + 2\sqrt{2}\gamma\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t}\right) \\ &\leq 2d\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^2t^2}{6bt + 2\sqrt{2}\gamma\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t}\right) \\ &= 2d\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})^2t}{6b + 2\sqrt{2}\gamma\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}\right).\end{aligned} \tag{62}$$

Bounding the probability of the second event Let $\mathcal{C} \in \mathcal{P}$ be a cluster, and let us denote $\boldsymbol{e}_m$ the m^{th} canonical vector of $\mathbb{R}^{|\mathcal{C}|}$. We have

$$\begin{aligned}
\left\|\overline{\boldsymbol{\Delta}^2}_{\mathcal{C}}(t)\right\| &= \frac{1}{|\mathcal{C}|}\left\|\sum_{m\in\mathcal{C}}\left(\sum_{\tau\in\mathcal{T}_m(t)} \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]\right)^2\right\| \\
&= \frac{1}{|\mathcal{C}|}\left\|\sum_{m\in\mathcal{C}} \boldsymbol{e}_m^\top \otimes \left(\sum_{\tau\in\mathcal{T}_m(t)} \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]\right)\right\|^2 \\
&= \frac{1}{|\mathcal{C}|}\left\|\sum_{\tau\in\bigcup_{m\in\mathcal{C}}\mathcal{T}_m(t)} \boldsymbol{e}_{m(\tau)}^\top \otimes \left(\boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]\right)\right\|^2 \\
&= \frac{1}{|\mathcal{C}|}\left\|\sum_{\tau\in\bigcup_{m\in\mathcal{C}}\mathcal{T}_m(t)} \boldsymbol{e}_{m(\tau)}^\top \otimes \boldsymbol{x}\boldsymbol{x}(\tau) - \mathbb{E}\left[\boldsymbol{e}_{m(\tau)} \otimes \boldsymbol{x}\boldsymbol{x}(\tau)|\mathcal{F}_{\tau-1}\right]\right\|^2,
\end{aligned}$$

where the last equality holds since $m(\tau)$ is measurable w.r.t. $\mathcal{F}_{\tau-1}$. We will apply the Corollary 2 to the set of time steps $\bigcup_{m\in\mathcal{C}}\mathcal{T}_m(t)$ and the adapted sequence $\boldsymbol{e}_{m(\tau)}^\top \otimes \boldsymbol{x}\boldsymbol{x}(\tau)$ of matrices in $\mathbb{R}^{d\times d|\mathcal{C}|}$. Hence we have

$$\begin{aligned}
&\mathbb{P}\left[\sqrt{\left\|\overline{\boldsymbol{\Delta}^2}_{\mathcal{C}}(t)\right\|} > \gamma t \min_{\mathcal{C}\in\mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})} \,\Big|\, \max_{m\in\mathcal{V}} |\mathcal{T}_m(t)| \le bt\right] \\
&\le d(1+|\mathcal{C}|)\exp\left(\frac{-3\gamma^2|\mathcal{C}|\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t^2}{6\sum_{m\in\mathcal{C}}|\mathcal{T}_m(t)| + 2\sqrt{2}\gamma\sqrt{|\mathcal{C}|\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}t}\right) \\
&\le d(1+|\mathcal{C}|)\exp\left(\frac{-3\gamma^2|\mathcal{C}|\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t}{6|\mathcal{C}|b + 2\sqrt{2}\gamma\sqrt{|\mathcal{C}|\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}}\right) \\
&= d(1+|\mathcal{C}|)\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t}{6b + 2\sqrt{2}\gamma\sqrt{\frac{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}{|\mathcal{C}|}}}\right) \\
&\le d(1+|\mathcal{C}|)\exp\left(\frac{-3\gamma^2\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})t}{6b + 2\sqrt{2}\gamma\sqrt{\frac{\min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C})}{\min_{\mathcal{C}\in\mathcal{P}}|\mathcal{C}|}}}\right)
\end{aligned} \tag{63}$$

Union bound We conclude the result of the statement via a union bound using Equation (60). □

Proposition 9 (Concentration of the empirical multi-task Gram matrix around the adapted one, simplified)**.** *Let $t \ge 1$, $b > 0$. Assume that $\max_{m\in\mathcal{V}}|\mathcal{T}_m(t)| \le bt$. Then we have:*

$$\mathbb{P}\left[\left\|\frac{\boldsymbol{A}_{\mathcal{V}}}{t} - \boldsymbol{V}_{\mathcal{V}}\right\|_{\text{op,RE}} > \gamma\right] \le 6d|\mathcal{V}|\exp\left(\frac{-3\gamma^2(\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}_{\mathcal{G}}(\mathcal{C})^2)t}{6b + 2\sqrt{2}\gamma}\right),$$

where $\tilde{c}_{\mathcal{G}}(\mathcal{C}) \coloneqq c_{\mathcal{G}}(\mathcal{C}) \wedge |\mathcal{C}| \quad \forall \mathcal{C} \in \mathcal{P}$.

Proof. The proof will rely on simple calculus inequalities. Hence, let $u = \min_{\mathcal{C}\in\mathcal{P}} c_{\mathcal{G}}(\mathcal{C}), v = \min_{\mathcal{C}\in\mathcal{P}} |\mathcal{C}|, f = 3\gamma^2, g = 6b, h = 2\sqrt{2}\gamma$, which are all positive. Then, we have

$$A_1 = \frac{fu}{f+g} \geq \frac{(u\wedge v)f}{f+g} \geq (u\wedge v)\frac{(1\wedge u\wedge v)f}{f+g(1\wedge u\wedge v)}$$
$$A_2 = \frac{fv}{f+g\frac{v}{u}} \geq \frac{(v\wedge u)f}{f+g\frac{v\wedge u}{u}} \geq \frac{(v\wedge u)f}{f+g} \geq (u\wedge v)\frac{(1\wedge u\wedge v)f}{f+(1\wedge u\wedge v)g}$$
$$A_3 = \frac{fv^2}{f+gv} \geq \frac{(v\wedge u)^2}{f+(v\wedge u)g} \geq (u\wedge v)\frac{(1\wedge u\wedge v)f}{f+(1\wedge u\wedge v)g}$$

where we used the fact that functions of the form $x \mapsto \frac{x}{\beta_1 x+\beta_2}$ for positive β_1, β_2 are increasing on $\mathbb{R}_+$.

As a final step, we use the inequality $\dfrac{(1\wedge x)f}{f+(1\wedge x)g} \geq \dfrac{x\wedge 1}{f+g}$ taken for $x = u\wedge v$, we apply the $\exp(-\cdot t)$ function and we use the result of Proposition 8, we deduce the result. □

B.3.3 From the true to the adapted Gram matrix

For all of the proofs in this subsection, we follow an approach similar to that of Oh et al. [2021]. In particular, we use their Lemma 10.

Theorem 7 (Lemma 10 of Oh et al. [2021])**.** *Under Assumption 2 on the context generating distribution, let* $t \geq 1$. *We have for any* $\boldsymbol{\theta} \in \mathbb{R}^d$*:*

$$\sum_{\boldsymbol{x}\in\mathcal{A}(t)} \mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}^\top \mathbb{1}\left\{\boldsymbol{x} \in \arg\max_{\tilde{\boldsymbol{x}}\in\mathcal{A}(t)} \langle\boldsymbol{\theta}, \tilde{\boldsymbol{x}}\rangle\right\}\right] \succcurlyeq \frac{1}{2\nu\omega}\overline{\boldsymbol{\Sigma}} \tag{64}$$

Proposition 10 (RE condition from the true to the adapted Gram matrix)**.** *Under Assumption 2, for any* $t \geq 1$*, the adapted Gram matrix* $\boldsymbol{V}_{\mathcal{V}}(t)$ *verifies the compatibility condition with constants* κ *and* $\dfrac{\phi}{\sqrt{2\nu\omega}}$.

Proof. For $t \geq 1$, we have

$$\mathbb{E}\left[\boldsymbol{x}(t)\boldsymbol{x}(t)^\top|\mathcal{F}_{t-1}\right] = \mathbb{E}\left[\sum_{\boldsymbol{x}\in\mathcal{A}(t)} \boldsymbol{x}(t)\boldsymbol{x}(t)^\top|\mathcal{F}_{t-1}\right] \tag{65}$$

Let $m \in \mathcal{V}$. We have

$$\begin{aligned}
\boldsymbol{V}_m(t) &= \frac{1}{t}\sum_{\tau\in\mathcal{T}_m(t)} \mathbb{E}\left[\boldsymbol{x}(\tau)\boldsymbol{x}(\tau)^\top|\mathcal{F}_{\tau-1}\right] \\
&= \frac{1}{t}\sum_{\tau\in\mathcal{T}_m(t)} \mathbb{E}\left[\mathbb{E}\left[\boldsymbol{x}(\tau)\boldsymbol{x}(\tau)^\top|\boldsymbol{\theta}_m(\tau-1), \mathcal{F}_{\tau-1}\right]|\mathcal{F}_{\tau-1}\right] \quad \text{(law of total expectation)} \\
&= \frac{1}{t}\sum_{\tau\in\mathcal{T}_m(t)} \mathbb{E}\left[\boldsymbol{x}(\tau)\boldsymbol{x}(\tau)^\top|\boldsymbol{\theta}_m(\tau-1)\right] \quad (\boldsymbol{x}(\tau) \text{ is fully determined by } \boldsymbol{\theta}_m(\tau-1)) \\
&= \frac{1}{t}\sum_{\tau\in\mathcal{T}_m(t)} \mathbb{E}\left[\sum_{\boldsymbol{x}\in\mathcal{A}(\tau)} \boldsymbol{x}\boldsymbol{x}^\top \mathbb{1}\left\{\boldsymbol{x} \in \arg\max_{\tilde{\boldsymbol{x}}\in\mathcal{A}(t)} \langle\boldsymbol{\theta}, \tilde{\boldsymbol{x}}\rangle\right\}|\boldsymbol{\theta}_m(\tau-1)\right] \\
&\succcurlyeq \frac{1}{2\nu\omega}\overline{\boldsymbol{\Sigma}} \quad \text{(by Theorem 7)}.
\end{aligned} \tag{66}$$

Now, let $\boldsymbol{Z} \in \mathcal{S}$, where $\mathcal{S}$ is defined with constant κ of Assumption 4. Then

$$\begin{aligned}
\sum_{m\in\mathcal{V}} \|\boldsymbol{z}\|_{\boldsymbol{V}_m(t)} &\geq \frac{1}{2\nu\omega}\sum_{m\in\mathcal{V}} \|\boldsymbol{z}_m\|_{\overline{\boldsymbol{\Sigma}}} \quad \text{by Equation (66)} \\
&\geq \frac{\phi^2}{2\nu\omega}\|\boldsymbol{Z}\|^2_{\text{RE}} \quad \text{(by Assumption 4)},
\end{aligned}$$

which finishes the proof. □

Theorem 8 (RE condition holding for the empirical multi-task Gram matrix, generalization of Theorem 2)**.** *Under assumptions 2 and 4, let $t \geq 1$, and let κ, ϕ be the constants from Assumption 4. Assume that $\max_{m \in \mathcal{V}} |\mathcal{T}_m(t)| \leq bt$. Then, for any $\gamma \in \left(0, \left(1 + \frac{a_2 + (1-\kappa)^+ \sqrt{2} w(\partial \mathcal{P})}{a_1}\right)^{-2}\right)$, the empirical multi-task Gram matrix verifies the RE condition with constants κ and $\hat{\phi}$, with*

$$\hat{\phi} = \tilde{\phi} \sqrt{1 - \gamma \left(1 + \frac{a_2 + (1-\kappa)^+ \sqrt{2} w(\partial \mathcal{P})}{a_1}\right)^2}, \tag{67}$$

with a probability at least equal to $1 - 6d|\mathcal{V}| \exp\left(\frac{-3\gamma^2 \tilde{\phi}^4 (\min_{\mathcal{C} \in \mathcal{P}} (\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}_{\mathcal{G}}(\mathcal{C})^2) t}{6b + 2\sqrt{2}\gamma \tilde{\phi}^2}\right)$, where $\tilde{\phi} \coloneqq \frac{\phi}{\sqrt{2\nu\omega}}$ and $\tilde{c}_{\mathcal{G}}(\mathcal{C}) \coloneqq c_{\mathcal{G}}(\mathcal{C}) \wedge |\mathcal{C}| \quad \forall \mathcal{C} \in \mathcal{P}$.

Proof. For the sake of readability, let $\tilde{\phi} = \frac{\phi}{\sqrt{2\nu\omega}}$ the compatibility constant of the adapted Gram matrix, according to Proposition 10. Then:

$$1 - 6d|\mathcal{V}| \exp\left(\frac{-3\gamma^2 \tilde{\phi}^4 (\min_{\mathcal{C} \in \mathcal{P}} (\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}_{\mathcal{G}}(\mathcal{C})^2) t}{6b + 2\sqrt{2}\gamma \tilde{\phi}^2}\right) \tag{68}$$

$$\leq \mathbb{P}\left[\left\| \frac{\boldsymbol{A}_{\mathcal{V}}}{t} - \boldsymbol{V}_{\mathcal{V}} \right\|_{\text{op,RE}} \leq \gamma \tilde{\phi}^2\right] \quad \text{(by Proposition 9)} \tag{69}$$

$$\leq \mathbb{P}\left[\frac{\boldsymbol{A}_{\mathcal{V}}}{t} \text{ satisfies the RE condition with constant } \kappa \text{ and } \hat{\phi}\right] \quad \text{(by Proposition 7)}, \tag{70}$$

where $\hat{\phi} = \tilde{\phi} \sqrt{1 - \gamma \left(1 + \frac{a_2 + (1-\kappa)^+ \sqrt{2} w(\partial \mathcal{P})}{a_1}\right)^2}$. ☐

B.4 Regret bound

Lemma 5 (Concentration of the fraction of observations per task)**.** *Assume that $|\mathcal{V}| \geq 2$. Then for $\delta \in (0, 1)$, we have with a probability at least $1 - \delta$:*

$$\max_{m \in \mathcal{V}} \frac{|\mathcal{T}_m(t)|}{t} \leq \frac{1}{|\mathcal{V}|} + 2\sqrt{\frac{1}{t|\mathcal{V}|} \log \frac{|\mathcal{V}|}{\delta}} + \frac{4}{3t} \log \frac{|\mathcal{V}|}{\delta}. \tag{71}$$

Proof. We have $|\mathcal{T}_m(t)| \coloneqq \sum_{\tau=1}^{t} [m(\tau) = m]$, where $\forall t, \forall m \in \mathcal{V}, \mathbb{P}[m(t) = m] = \frac{1}{|\mathcal{V}|}$, meaning that the binary variable $[m(t) = m]$ follows a Bernoulli distribution $\mathcal{B}(\frac{1}{\mathcal{V}})$. Then, the random variable $X_t \coloneqq [m(t) = m] - \frac{1}{|\mathcal{V}|}$ has mean 0, variance $\frac{1}{|\mathcal{V}|}(1 - \frac{1}{|\mathcal{V}|})$, and verifies $|X_t| \leq 1 - \frac{1}{|\mathcal{V}|}$ since $|\mathcal{V}| \geq 2$. As a result, via the Bernstein inequality, we have for any $m \in \mathcal{V}$, and for any $w \geq 0$,

$$\mathbb{P}\left[\frac{|\mathcal{T}_m(t)|}{t} \geq \frac{1}{|\mathcal{V}|} + w\right] \leq \exp\left(-\frac{tw^2}{2(1 - \frac{1}{|\mathcal{V}|})(\frac{1}{|\mathcal{V}|} + \frac{w}{3})}\right) \leq \exp\left(-\frac{tw^2}{2(\frac{1}{|\mathcal{V}|} + \frac{w}{3})}\right)$$

For the right-hand side to hold with a probability at most $\delta \in (0, 1)$, it is sufficient to have

$$t \frac{w^2}{2(\frac{1}{|\mathcal{V}|} + \frac{w}{3})} \geq \log \frac{1}{\delta}$$

$$\Longleftarrow \frac{w^2}{2} \geq \frac{2 \frac{1}{|\mathcal{V}|} \log \frac{1}{\delta}}{t} \text{ and } \frac{w^2}{2} \geq \frac{2w \log \frac{1}{\delta}}{3t}$$

$$\Longleftarrow w = 2\sqrt{\frac{\frac{1}{|\mathcal{V}|} \log \frac{1}{\delta}}{t}} + \frac{4 \log \frac{1}{\delta}}{3t}$$

Hence, and via a union bound, we get

$$\mathbb{P}\left[\frac{|\mathcal{T}_m(t)|}{t} \geq \frac{1}{|\mathcal{V}|} + 2\sqrt{\frac{1}{|\mathcal{V}|}\log\frac{1}{\delta}} + \frac{4}{3t}\log\frac{1}{\delta}\right] \leq \delta$$

$$\Longrightarrow \mathbb{P}\left[\max_{m\in\mathcal{V}}\frac{|\mathcal{T}_m(t)|}{t} \geq \frac{1}{|\mathcal{V}|} + 2\sqrt{\frac{\frac{1}{|\mathcal{V}|}\log\frac{1}{\delta}}{t}} + \frac{4\log\frac{1}{\delta}}{3t}\right] \leq |\mathcal{V}|\delta$$

The result is obtained by adjusting the value of δ. □

Theorem 9 (Regret bound, generalization of Theorem 3)**.** *Let the mean horizon per node be* $\overline{T} = \frac{T}{|\mathcal{V}|}$*. Under assumptions1 to 4 and* $\kappa > 0$*, the expected regret of the Network Lasso Bandit algorithm is upper bounded as follows:*

$$\mathcal{R}(\overline{T}) = \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G},\mathbf{\Theta},\alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right) + \frac{1}{A}\log(d|\mathcal{V}|) + \sqrt{|\mathcal{V}|}\right),$$

with $A = \dfrac{3\gamma^2\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}^2_{\mathcal{G}}(\mathcal{C}))}{6\frac{\log(|\mathcal{V}|)}{\sqrt{|\mathcal{V}|}} + 2\sqrt{2}\gamma}$.

Proof. For any time step t, we will define a list of good events under which the Oracle inequality and the RE condition for the empirical multi-task Gram matrix both hold with high probability. Then, we will use those bounds to sum up over time steps until horizon T.

Good events We formalize these requirements as three families of time-depending "good" events.

- $G_{\text{pro}}(t)$ is the event that the mean of the empirical process bounded by $\alpha(t)$ up to a constant c, which is equivalent to saying that it converges:

$$G_{\text{pro}}(t) \coloneqq \left\{\frac{1}{t}\|\boldsymbol{K}\|_F \leq \frac{\alpha(t)}{\alpha_0}\right\} \tag{72}$$

- $G_{\text{sel}}(t)$ is the event that the number of selections of all tasks is bounded by its expected value up to a small constant $\rho(t)$

$$G_{\text{sel}}(t) \coloneqq \left\{\max_{m\in\mathcal{V}}\frac{|\mathcal{T}_m(t)|}{t} \leq \frac{1}{|\mathcal{V}|} + \frac{\rho(t)}{t}\right\} \tag{73}$$

- $G_{\text{RE}}(t)$ is the event that the empirical multi-task Gram matrix $\frac{1}{t}\boldsymbol{A}_{\mathcal{V}}(t)$ satisfies the RE condition.

$$G_{\text{RE}}(t) \coloneqq \left\{\frac{1}{t}\boldsymbol{A}_{\mathcal{V}}(t) \text{ verifies the RE condition with constants } \kappa, \hat{\phi}\right\} \tag{74}$$

Event $G_{\text{pro}}(t)$ is the most straightforward to cover since our bound on the empirical process given in Lemma 3 holds with a probability of at least $1-\delta(t)$, thus:

$$\mathbb{P}\left[G_{\text{pro}}(t)^c | G_{\text{sel}}(t)\right] \leq \delta(t), \tag{75}$$

where we included the time dependency on $\delta(t)$ in contrast to the previous section. This way we emphasize to adjust $\delta(t)$ after each round, to guarantee a sub linear regret bound. The probability of event $G_{\text{sel}}(t)$ can be determined using Bernstein's inequality:

From Lemma 5 we can select $\rho(t) = 2\sqrt{\frac{t}{|\mathcal{V}|}\log\frac{|\mathcal{V}|}{\delta_{\text{sel}}(t)}} + \frac{4}{3}\log\frac{|\mathcal{V}|}{\delta_{\text{sel}}(t)}$ as well as $\mathbb{P}\left[G_{\text{sel}}(t)^c\right] \leq \delta_{\text{sel}}(t)$.

B.4.1 Instantaneous regret decomposition

Now, given the event probabilities, we condition the instantaneous regret $r(t)$ on the good events at a time $t > t_0$. We have for its expectation:

$$\begin{aligned}\mathbb{E}\left[r(t)\right] &\leq \mathbb{E}\left[r(t)|G_{\text{sel}}(t)\right] + 2\mathbb{P}\left[G_{\text{sel}}(t)^c\right] \\ &\leq \mathbb{E}\left[r(t)|G_{\text{pro}}(t) \cap G_{\text{RE}}(t) \cap G_{\text{sel}}(t)\right] \\ &+ 2\left(\mathbb{P}\left[G_{\text{pro}}(t)^c|G_{\text{sel}}(t)\right] + \mathbb{P}\left[G_{\text{RE}}(t)^c|G_{\text{sel}}(t)\right] + \mathbb{P}\left[G_{\text{sel}}(t)^c\right]\right),\end{aligned} \tag{76}$$

where we used the worst case bound $r(t) \leq 2$ if any one of the good events does not hold.

Bounding the regret Inserting our results of the event probabilities, the oracle inequality and the decomposition of the expected instantaneous regret in Equation (76) and bounding the sum over rounds, yields the final result. Thus, we start by bounding the sum over the first term i.e. the expected regret in case all good events hold:

$$\sum_{t=1}^{T} \mathbb{E}\left[r(t)|G_{\text{pro}}(t) \cap G_{\text{RE}}(t) \cap G_{\text{sel}}(t)\right] \leq \sum_{t=1}^{T} \left\|\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}(t)\right\|_F$$

Taking the result of our oracle inequality in Theorem 5, we point out that only $\alpha(t)$ is time dependent such that the rest of the terms can be pulled outside the sum:

$$\begin{aligned}\sum_{t=1}^{T} \left\|\boldsymbol{\Theta} - \hat{\boldsymbol{\Theta}}(t)\right\|_F &\leq \sum_{t=1}^{T} 2\frac{\alpha_0 \sigma}{\hat{\phi}^2 \sqrt{t}} f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \sqrt{1 + 2b\sqrt{|\mathcal{V}| \log \frac{1}{\delta(t)}} + 2b \log \frac{1}{\delta(t)}} \\ &= \frac{2\alpha_0 \sigma}{\hat{\phi}^2} f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \sum_{t=1}^{T} \sqrt{\frac{1}{t} + \frac{2b}{t}\sqrt{2|\mathcal{V}| \log(t)} + \frac{4b}{t} \log(t)} \\ &\leq \frac{2\alpha_0 \sigma}{\hat{\phi}^2} f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \int_0^T \frac{1}{\sqrt{t}} + \sqrt{\frac{2b}{t}\left(\sqrt{2|\mathcal{V}| \log(T)} + 2\log(T)\right)}\, dt \\ &\leq \frac{2\alpha_0 \sigma}{\hat{\phi}^2} f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \Bigg(2\sqrt{T} + \left(\frac{\sqrt{8T}}{|\mathcal{V}|} + 4\sqrt[4]{\frac{32 \log(|\mathcal{V}|T) T}{|\mathcal{V}|}} + \sqrt{\frac{16}{3} \log(|\mathcal{V}|T) \log(T)}\right) \\ &\left(\sqrt[4]{2|\mathcal{V}| \log(T)} + \sqrt{2\log(T)}\right)\Bigg) \\ &= \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2} \left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}| \log(\overline{T}|\mathcal{V}|)}\right)\right),\end{aligned}$$

where

$$f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) := \left(a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) + \sqrt{2}\mathbb{1}_{\leq 1}(\kappa) w(\partial\mathcal{P})\right)\left(\frac{a_2(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) + \sqrt{2}\mathbb{1}_{\leq 1}(\kappa) w(\partial\mathcal{P})}{a_1(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0) \min\limits_{\mathcal{C} \in \mathcal{P}} \sqrt{c_{\mathcal{G}}(\mathcal{C})}} + 1\right).$$

We upper bounded the sum with an integral i.e. $\sum_{t=1}^{T} g(t) \leq \int_0^T g(t) dt$ for monotonically decreasing functions $g(t)$ in the last inequality. Also b is the bound on the concentration of the fraction of observation per task provided by Lemma 5. For $t_0 = \sqrt{|\mathcal{V}|}$ we find by inserting the result to Lemma 5 for all $t > t_0$:

$$
\begin{aligned}
\frac{1}{|\mathcal{V}|} + 2\sqrt{\frac{1}{t|\mathcal{V}|}\log\frac{|\mathcal{V}|}{\delta}} + \frac{4}{3t}\log\frac{|\mathcal{V}|}{\delta} &\leq \frac{1}{|\mathcal{V}|} + 2\sqrt{\frac{2\log\left(|\mathcal{V}|\sqrt{|\mathcal{V}|}\right)}{\sqrt{|\mathcal{V}|}|\mathcal{V}|}} + \frac{8\log\left(|\mathcal{V}|\sqrt{|\mathcal{V}|}\right)}{3\sqrt{|\mathcal{V}|}} \\
&= \frac{1}{|\mathcal{V}|} + \frac{2}{\sqrt{|\mathcal{V}|}}\left[\sqrt{\frac{3}{\sqrt{|\mathcal{V}|}}\log(|\mathcal{V}|)} + 2\log(|\mathcal{V}|)\right] \\
&= \mathcal{O}\left(\frac{\log(|\mathcal{V}|)}{\sqrt{|\mathcal{V}|}}\right) = b.
\end{aligned}
$$

Finally we bound the sum over the instantaneous regret term for the bad events:

$$
\sum_{t=1}^{T} 2\left(\mathbb{P}\left[G_{\text{pro}}(t)^c | G_{\text{sel}}(t)\right] + \mathbb{P}\left[G_{\text{RE}}(t)^c | G_{\text{sel}}(t)\right] + \mathbb{P}\left[G_{\text{sel}}(t)^c\right]\right)
$$

By construction, we have $\max(\mathbb{P}\left[G_{\text{pro}}(t)^c | G_{\text{sel}}(t)\right], \mathbb{P}\left[G_{\text{sel}}(t)^c\right]) \leq \delta(t) = \frac{1}{t^2}$. Hence,

$$
\sum_{t=1}^{T} \mathbb{P}\left[G_{\text{pro}}(t)^c | G_{\text{sel}}(t)\right] + \mathbb{P}\left[G_{\text{sel}}(t)^c\right] \leq 2\sum_{t=1}^{T}\frac{1}{t^2} \leq 2\left(1 + \int_1^T \frac{dt}{t^2}\right) \leq 4 \tag{77}
$$

As for the RE condition event, letting $A := \dfrac{3\gamma^2 \min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C}) \wedge \tilde{c}_{\mathcal{G}}^2(\mathcal{C}))}{6b + 2\sqrt{2}\gamma}$, we have for any $t_0 \geq 1$

$$
\begin{aligned}
\sum_{t=t_0}^{T} \mathbb{P}\left[G_{\text{RE}}(t)^c | G_{\text{sel}}(t)\right] &\leq 6d|\mathcal{V}|\sum_{t=t_0}^{T}\exp(-At) \quad \text{(by Theorem 8)} \\
&\leq 6d|\mathcal{V}|\frac{e^{-At_0}}{1 - e^{-A}} \leq 6d|\mathcal{V}|e^{-At_0}\left(1 + \frac{1}{A}\right) \\
&\leq 6d|\mathcal{V}|e^{-At_0}\left(1 + \frac{1}{A}\right)
\end{aligned}
$$

where in the last line, we used the inequality $\exp(A) \geq A + 1$. Hence, for any $u > 0$, choosing

$$
t_0 = \left\lceil \sqrt{|\mathcal{V}|} \right\rceil \vee \left\lceil \frac{1}{A}\log\left(\frac{6d|\mathcal{V}|(1 + \frac{1}{A})}{u}\right) \right\rceil
$$

implies that $\sum_{t=t_0}^{T} \mathbb{P}\left[G_{\text{RE}}(t)^c | G_{\text{sel}}(t)\right] \leq u$. Now, we simply have to insert all our results into the sum of instantaneous regrets:

$$
\begin{aligned}
\mathcal{R}(\overline{T}) &\leq t_0 + 2u + 8 + \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right) \\
&\leq \left\lceil \sqrt{|\mathcal{V}|} \right\rceil + \left\lceil \frac{1}{A}\log\left(\frac{6d|\mathcal{V}|(1+\frac{1}{A})}{u}\right)\right\rceil + 2u + 8 \\
&+ \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right) \\
&\leq \left\lceil \sqrt{|\mathcal{V}|} \right\rceil + \left\lceil \frac{1}{A}\log(12d|\mathcal{V}|(1+A))\right\rceil + \frac{1}{A} + 8 \\
&+ \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right) \\
&\leq \left\lceil \sqrt{|\mathcal{V}|} \right\rceil + \left\lceil \frac{1}{A}\log(12d|\mathcal{V}|(1+A))\right\rceil + \frac{1}{A} + 8 \\
&+ \mathcal{O}\left(\frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right) \\
&= \mathcal{O}\left(\sqrt{|\mathcal{V}|} + \frac{1}{A}\log(d|\mathcal{V}|) + \frac{\alpha_0 f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\hat{\phi}^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right) \\
&= \mathcal{O}\left(\sqrt{|\mathcal{V}|} + \frac{1}{A}\log(d|\mathcal{V}|) + \frac{\alpha_0 \nu\omega f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)\sqrt{\overline{T}}}{\phi^2}\left(\sqrt{|\mathcal{V}|} + \sqrt{\log(\overline{T}|\mathcal{V}|)} + \sqrt[4]{|\mathcal{V}|\log(\overline{T}|\mathcal{V}|)}\right)\right),
\end{aligned}
$$

where we set $u = \frac{1}{2A}$ in the third inequality.

□

Proof of Corollary 1. Assuming $\frac{w(\partial\mathcal{P})(\psi+2\kappa)}{\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}} \leq \Omega$, with some positive constant $\Omega < 1$ and setting $\alpha_0 = \frac{1}{\psi w(\partial\mathcal{P})}$ then the term $f(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0)$ can be bounded as:

$$
\begin{aligned}
f\left(\mathcal{G}, \boldsymbol{\Theta}, \alpha_0 = \frac{1}{\psi w(\partial\mathcal{P})}\right) &= a_2\left(\frac{a_2}{a_1 \min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}} + 1\right) \\
&= \left(w(\partial\mathcal{P})\left(\psi + \sqrt{2}\kappa \max_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\right)\right)\left(\frac{w(\partial\mathcal{P})\left(\psi + \sqrt{2}\kappa\max_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\right)}{\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})} - w(\partial\mathcal{P})(\psi+2\kappa)} + 1\right) \\
&= \mathcal{O}\left(\frac{w(\partial\mathcal{P})^2 \max_{\mathcal{C}\in\mathcal{P}}\iota_{\mathcal{G}}(\mathcal{C})}{(1-\Omega)\min_{\mathcal{C}\in\mathcal{P}}\sqrt{c_{\mathcal{G}}(\mathcal{C})}}\right).
\end{aligned}
$$

Ω acts as a threshold for the quality of any graph the satisfy this bound. Similarly we can find a bound on the term $\frac{1}{A}$:

$$
\begin{aligned}
\frac{1}{A} &= \frac{6\frac{\log(|\mathcal{V}|)}{\sqrt{|\mathcal{V}|}} + \frac{1}{\sqrt{2}}\left(1+\frac{a_2}{a_1}\right)^{-2}}{\frac{3}{4}\left(1+\frac{a_2}{a_1}\right)^{-4}\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C})\wedge\tilde{c}^2_{\mathcal{G}}(\mathcal{C}))} \\
&\le \frac{6\frac{\log(|\mathcal{V}|)}{\sqrt{|\mathcal{V}|}} + \frac{1}{\sqrt{2}}\left(1+\frac{w(\partial\mathcal{P})\left(\psi+\sqrt{2}\kappa\max\limits_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\right)}{1-\Omega}\right)^{-2}}{\frac{3}{4}\left(1+\frac{w(\partial\mathcal{P})(\psi+\sqrt{2}\kappa)\max\limits_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}}{1-\Omega}\right)^{-4}\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C})\wedge\tilde{c}^2_{\mathcal{G}}(\mathcal{C}))} \\
&= \mathcal{O}\left(\frac{\left(1+\frac{w(\partial\mathcal{P})\left(\psi+\sqrt{2}\kappa\max\limits_{\mathcal{C}\in\mathcal{P}}\sqrt{\iota_{\mathcal{G}}(\mathcal{C})}\right)}{1-\Omega}\right)^{2}}{\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C})\wedge\tilde{c}^2_{\mathcal{G}}(\mathcal{C}))}\right) \\
&= \mathcal{O}\left(\frac{w(\partial\mathcal{P})^2\max\limits_{\mathcal{C}\in\mathcal{P}}\iota_{\mathcal{G}}(\mathcal{C})}{(1-\Omega)^2\min_{\mathcal{C}\in\mathcal{P}}(\tilde{c}_{\mathcal{G}}(\mathcal{C})\wedge\tilde{c}^2_{\mathcal{G}}(\mathcal{C}))}\right)
\end{aligned}
$$

Inserting the terms into the regret bound yields the final result.

□

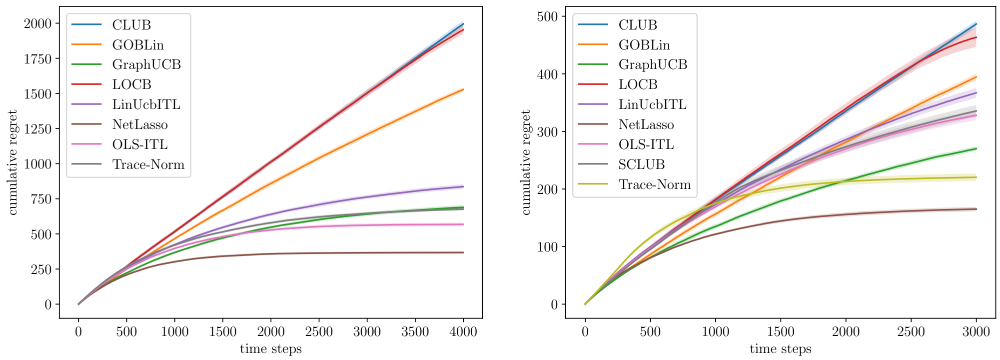
C Additional experimental details

C.1 About experiments of the main paper

The experiments have been conducted with an intel i7 CPU with 12 2.6 GHz cores and 32 GB of RAM. The two experiments with the highest number of tasks (200) and dimension (80) take about 8 hours, parallelized over the 12 cores.

To generate clusters, we generate $|\mathcal{P}|$ variables $v_{i i\in\mathcal{P}}$ from the uniform distribution, then we use them to construct a categorical distribution with probabilities proportional to e^{v_i}. These probabilities defines the cluster proportions.
We included the algorithm of [Cella et al., 2023] as a baseline for the experiments to cover the particular case that the number of clusters is lower than the number of dimensions. Indeed, the cluster structure of $\mathbf{\Theta}$ can be mathematically written as $\mathbf{\Theta} = \sum_{C\in\mathcal{P}} \mathbf{1}_C\boldsymbol{\theta}_C^\top$, where $\mathbf{1}_C$ is the indicator vector of cluster C (coordinates equal to 1 on the nodes belonging to C and zeros elsewhere) and $\boldsymbol{\theta}_C$ is the true vector of every node in C. The range of $\mathbf{\Theta}$ is equal to the span of $\mathbf{1}_C; C\in\mathcal{P}$, implying that its rank is at most equal to $\min(d, |\mathcal{P}|)$. It will then satisfy the low-rank assumption for $|\mathcal{P}| < d$.

Our results clearly demonstrate an improvement compared to the other baselines. Our policy performs significantly better than the rest beyond the error margins, covering one standard deviation at ten repetitions. We provide results for up to $|\mathcal{V}| = 200$ nodes showing the effective transfer of knowledge between nodes within the graph.



(a) $|\mathcal{V}| = 200, |\mathcal{P}| = 25, d = 10, p = 0.5, q = 0.05$ (b) $|\mathcal{V}| = 50, |\mathcal{P}| = 5, d = 80, p = 0.8, q = 0.2$

Figure 1: Synthetic data experiments showing the cumulative regret of Network Lasso Policy as a function of time-steps compared to other baselines, for different choices of $|\mathcal{V}|, |\mathcal{P}|, d, p$ and q.

C.2 Solving the Network Lasso problem

We implement the Primal-Dual algorithm proposed in Jung [2020] to solve the Network Lasso problem but we do not vectorize the matrices (in the sense of stacking their columns into a vector), which speeds up computation.

C.3 Algebraic connectivity vs topological centrality index

Given two fully connected graphs weightless $\mathcal{G}_1$ and $\mathcal{G}_2$ with size 100 each, we progressively link them by edges and construct the Laplacian $\boldsymbol{L}$ of the resulting graph $\mathcal{G}$. We measure the minimum topological centrality index $\min_{1 \leq i \in 200} (\boldsymbol{L}_{\mathcal{C}}^{\dagger})_{ii}^{-1}$, and the algebraic connectivity, i.e. the minimum non-null eigenvalue of $\boldsymbol{L}$.

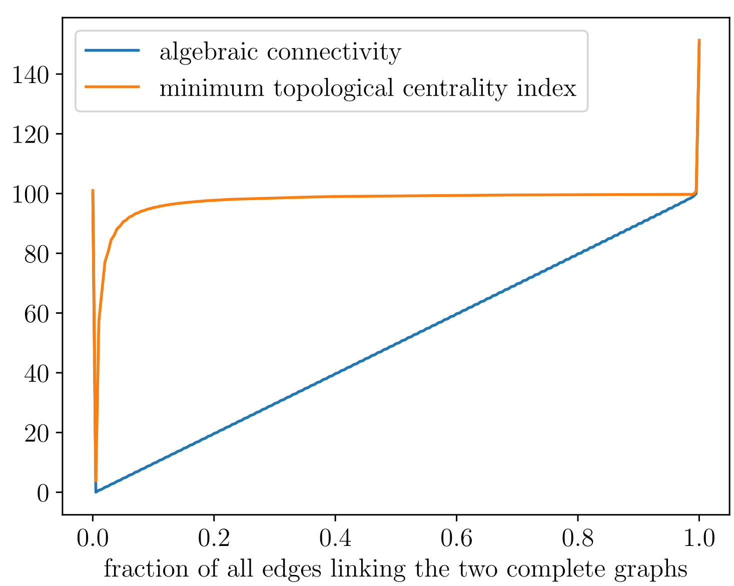


Figure 2: Minimum Topological centrality index vs Algebraic Connectivity, for a graph formed by connecting two fully connected initial graphs $\mathcal{G}_1, \mathcal{G}_2$ with size 100 each.

D Technical Appendices and Supplementary Material

Technical appendices with additional results, figures, graphs and proofs may be submitted with the paper submission before the full submission deadline (see above), or as a separate PDF in the ZIP file below before the supplementary material deadline. There is no page limit for the technical appendices.