Multitask Online Learning: Listen to the Neighborhood Buzz

Juliette Achddou
Università degli Studi di Milano
Italy

Nicolò Cesa-Bianchi
Università degli Studi di Milano
and Politecnico di Milano, Italy

Pierre Laforgue
Università degli Studi di Milano
Italy

Abstract

We study multitask online learning in a setting where agents can only exchange information with their neighbors on a given arbitrary communication network. We introduce $\texttt{MT-CO}_2\texttt{OL}$, a decentralized algorithm for this setting whose regret depends on the interplay between the task similarities and the network structure. Our analysis shows that the regret of $\texttt{MT-CO}_2\texttt{OL}$ is never worse (up to constants) than the bound obtained when agents do not share information. On the other hand, our bounds significantly improve when neighboring agents operate on similar tasks. In addition, we prove that our algorithm can be made differentially private with a negligible impact on the regret. Finally, we provide experimental support for our theory.

1 INTRODUCTION

Many real-world applications, including recommendation, personalized medicine, or environmental monitoring, require learning a personalized service offered to multiple clients. These problems are typically studied using multitask learning, or personalized federated learning when privacy is of concern. The key idea behind these techniques is that sharing information among similar clients may help learn faster. In this work we study multitask learning in an online convex optimization setting where multiple agents share information across a communication network to minimize their local regrets. We focus on decentralized algorithms, that operate without a central coordinating entity and only have a local knowledge of the communication network. Motivated by scenarios in which long-range communication is costly (e.g., in sensor networks) or slow (e.g., in advertising/financial networks, where data arrive at very high rates), we assume agents can only communicate with their neighbors in the network. Our regret analysis applies to the general multitask setting, where each agent is solving a potentially different online learning problem. In our decentralized environment, we do not require agents to work in a synchronized fashion. Rather, agents predict according to some unknown sequence of agent activations. We consider both deterministic (i.e., oblivious adversarial) or stochastic activation sequences.

It is well known that the optimal regret in single-agent single-task online convex optimization is $\mathcal{O}(\sqrt{T})$, achieved, for example, by the `FTRL` algorithm (Orabona, 2019). In the multi-agent setting with N agents, one can trivially achieve regret $\mathcal{O}\big(\sqrt{NT}\big)$ simply by running N independent instances of `FTRL` (no communication). Cesa-Bianchi et al. (2020) consider a multi-agent single-task setting where each active agent sends the current loss to their neighbors. In this setup, they show a regret bound in $\mathcal{O}\big(\sqrt{\alpha(G)\,T}\big)$, where $\alpha(G) \le N$ is the independence number of the communication graph G (unknown to the agents). They also show that when agents know G and active agents have access to the predictions of their neighbors, then the regret bound becomes $\mathcal{O}\big(\sqrt{\gamma(G)\,T}\big)$, where $\gamma(G) \le \alpha(G)$ is the domination number of G. The multi-agent multitask setting in online convex optimization was studied by Cesa-Bianchi et al. (2022) for the case when G is a clique. They prove a regret bound of order $\sqrt{1+\sigma(N-1)}\sqrt{T}$, where σ is the variance of the set of best local models for the tasks. This bound is achieved by a multitask variant of `FTRL` based on sharing the loss gradients.

In this work we consider the multitask setting when G is arbitrary, which—to the best of our knowledge—was never investigated in the context of online convex optimization. We introduce $\texttt{MT-CO}_2\texttt{OL}$, a new decentralized variant of `FTRL` where agents can only communicate with their neighbors. We show $\mathcal{O}\big(\sqrt{T}\big)$ regret bounds, for adversarial and stochastic activations, in which the scaling $\sqrt{1+\sigma(N-1)}$ for the clique is replaced by $\sqrt{1+\sigma_j(N_j-1)}$ summed over agents j, where σ_j is

Proceedings of the 27[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2024, Valencia, Spain. PMLR: Volume 238.

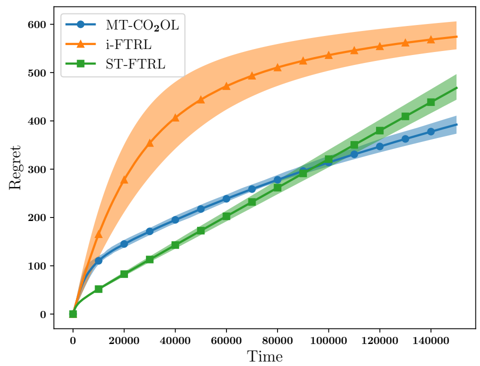


Figure 1: Multitask regret over time of `MT-CO`$_2$`OL` on a random communication graph with stochastic activations against two baselines: `i-FTRL` (N independent instances of `FTRL`) and `ST-FTRL` (the multi-agent single-task algorithm of Cesa-Bianchi et al. (2020)).

the task variance in the neighborhood of j of size N_j (Theorem 3). We also recover the previously known bounds for single-task and multitask settings as special cases of our result. Figure 1 shows that, when the task similarity is in the right range, `MT-CO`$_2$`OL` outperforms both multi-agent `FTRL` without communication among agents, and multi-agent single-task `FTRL`. See Section 5 for more details and experiments.

We also prove a lower bound (Theorem 6) showing that our regret upper bounds are tight in some important special cases, such as regular communication graphs. Finally, we show that `MT-CO`$_2$`OL` can be made differentially private, and prove that privacy only degrades the multitask regret by a term polylogarithmic in T (see Theorem 8). This allows us to identify a privacy threshold above which sharing information no longer benefits the agents.

1.1 Related works

We review related works by focusing on multi-agent online learning with adversarial (nonstochastic) losses. In this area, we distinguish three main threads of research: single-task (or cooperative) learning, multitask (or heterogeneous) learning, and distributed optimization. We also distinguish between synchronous (all agents are active in each round) and asynchronous (one agent is active in each round) activation models. To the best of our knowledge, the study of cooperative online learning was initiated by Awerbuch and Kleinberg (2008), who studied a cooperative-competitive synchronous bandit model in which agents are partitioned in unknown clusters and agents in the same cluster receive the same losses (i.e., solve a single-task problem) whereas the losses of agents in distinct clusters may be different.

Single-task. The synchronous bandit model was also studied by Cesa-Bianchi et al. (2019) and Bar-On and Mansour (2019) in a network of agents where communication delays (based on the shortest-path distance) are taken into account. Their analysis was extended to linear bandits by Ito et al. (2020), and to linear semi-bandits by Della Vecchia and Cesari (2021). Cesa-Bianchi et al. (2020) investigate the asynchronous online convex optimization setting without delays and where agents can only talk to their neighbors. Their analysis is extended to neighborhoods of higher order by Van der Hoeven et al. (2022), who also take communication delays into account. Hsieh et al. (2022) and Jiang et al. (2021) consider the same setting, but with a more abstract model of delays, not necessarily induced by shortest-path distance on a graph.

Multitask. One of the earliest contributions in multitask online learning is by Cavallanti et al. (2010), who introduce an asynchronous multitask version of the Perceptron algorithm for binary classification. Murugesan et al. (2016) extend this idea to a model where task similarity is also learned. Another extension is considered by Cesa-Bianchi et al. (2022), who introduce a multitask version of Online Mirror Descent with arbitrary regularizers. See also Saha et al. (2011); Zhang et al. (2018); Li et al. (2019) for additional multitask online algorithms without performance bounds. Recently, Herbster et al. (2021) investigated an asynchronous bandit model on a social network in which the network's partition induced by the labeling assigning the best local action to each node has a small resistance-weighted cutsize. Finally, note the work by Sinha and Vaze (2023), where the authors treat tasks as constraint generators (instead of loss generators), and achieve sublinear regret for online convex optimization.

Distributed online optimization. Yan et al. (2012) introduced a synchronous multitask setting where the comparator is the best global prediction for all tasks and the system's performance is measured according to $\max_{j\in[N]} \sum_{i=1}^{N} \sum_{t=1}^{T} \ell_t^i(x_t^j)$, where $\ell_t^i(x_t^j)$ is the loss at time t for the task of agent i evaluated on the prediction at time t of agent j. Crucially, at time t each agent j only observes $\nabla\ell_t^j(x_t^j)$, and communication is used to gather information about other tasks. Hosseini et al. (2013) extended the analysis from strongly convex losses to general convex losses. Yi and Vojnović (2023) consider the bandit case, in which each agent observes the loss incurred rather than the gradient. Note that the synchronous bandit model by Cesa-Bianchi et al. (2019) is a special case of this setting (when $\ell_t^1 = \ldots = \ell_t^N$ for all t). However, their regret bound scales with the independence number of the communication graph, as opposed to spectral quantities such as the inverse of the spectral gap in distributed online optimization.

2 PROBLEM SETTING

In this section, we introduce the multitask online learning setting, and recall important existing results.

Setting. Consider N learning agents located at the nodes of a communication network described by an undirected graph $G = (V, E)$, where $V = [N] \coloneqq \{1, \dots, N\}$. Let $\mathcal{N}_i = \{i\} \cup \{j \in [N] : (i,j) \in E\}$ be the neighbors of agent i in G (including i itself), and $N_i = |\mathcal{N}_i|$. We denote $N_{\min} = \min_{i \in V} N_i$, and $N_{\max} = \max_{i \in V} N_i$. We assume that agents ignore the graph, but know the identity of their neighbors. Let $\mathcal{X} \subset \mathbb{R}^d$ be the agents common decision space, and $\ell_1, \ell_2, \dots$ a sequence of convex losses from $\mathcal{X}$ to $\mathbb{R}$, secretly chosen by an oblivious adversary. The learning process is as follows. For $t = 1, 2, \dots$

1. some agent $i_t \in [N]$ is activated
2. i_t may *fetch* information from its neighbors $j \in \mathcal{N}_{i_t}$
3. i_t predicts $x_t \in \mathcal{X}$
4. i_t incurs the loss $\ell_t(x_t)$ and observes $g_t \in \partial \ell_t(x_t)$
5. i_t may *send* information to its neighbors $j \in \mathcal{N}_{i_t}$.

Communication is limited to steps 2 and 5, and only along edges incident on the active agent, hereby enforcing a decentralized learning process. As in distributed optimization algorithms (Hosseini et al., 2013), we use step 2 to fetch model-related information, and step 5 to send gradient-related information, see Sections 3 and 4. As shown in our analysis, both communication steps are key to obtain strong regret guarantees.

We measure performance through the *multitask regret*, defined as the sum of the agents' individual regrets. The individual regret measures the performance of a single agent against the best decision in hindsight for its *own personal task*, i.e., for the sequence of losses associated to the time steps it is active. Namely, agent i aims at minimizing for any $u^{(i)} \in \mathcal{X}$ its local regret $\sum_t \big(\ell_t(x_t) - \ell_t(u^{(i)})\big)\, \mathbb{I}\{i_t = i\}$, and the global objective is thus to control for any $u^{(1)}, \dots, u^{(N)} \in \mathcal{X}$ the multitask regret $\sum_{i=1}^N \sum_{t=1}^T \big(\ell_t(x_t) - \ell_t(u^{(i)})\big)\, \mathbb{I}\{i_t = i\}$. Equivalently, this amounts to minimizing

$$R_T(U) = \sum_{t=1}^T \ell_t(x_t) - \ell_t(U_{i_t:}) \tag{1}$$

for any horizon T and multitask comparator $U \in \mathcal{U}$, where $\mathcal{U} \coloneqq \big\{U \in \mathbb{R}^{N \times d} \colon U_{i:} \in \mathcal{X} \text{ for all } i\big\}$. Note that for simplicity in the rest of the paper we set $\mathcal{X} = \{x \in \mathbb{R}^d \colon \|x\|_2 \le 1\}$, and assume that all loss functions are 1-Lipschitz, i.e., we have $\|g_t\|_2 \le 1$ for all t. Our analysis can be readily extended to any ball of generic radius D and L-Lipschitz losses, up to scaling each regret bound by DL. As pointed out in the introduction, a naive approach to minimizing (1) consists in running N independent instances of FTRL, without making use of steps 2 and 5. By Jensen's inequality, such strategy satisfies for any $U \in \mathcal{U}$

$$R_T(U) = \sum_{i=1}^N 2\sqrt{T_i} \le 2\sqrt{NT}\,, \tag{2}$$

where $T_i = \sum_{t=1}^T \mathbb{I}\{i_t = i\}$. The purpose of this work is to introduce and analyze a new algorithm whose regret improves on (2) in terms of N. Our bounds should depend on the interplay between the task similarity and the structure of the communication graph.

Remark 1 (Comparison to Cesa-Bianchi et al. (2020))**.** *Although Cesa-Bianchi et al. (2020) also consider arbitrary communication graphs, their setting significantly differs from ours. Recall that they work in a single-task setting (where $u^{(1)} = u^{(2)} = \ldots = u^{(N)}$). Hence, their proof techniques are significantly different from ours. Moreover, while we only rely on gradient feedback, where the gradient is evaluated at the current prediction x_t, in their setting the learner is free to compute gradients at any point in the decision space. On the other hand, their communication model is restricted to sharing gradients while we can also fetch predictions.*

Preliminaries. Our algorithm builds upon MT-FTRL (Cesa-Bianchi et al., 2022), designed for the clique case. We recall important facts about this algorithm. MT-FTRL combines an instance of FTRL with a carefully chosen Mahalanobis regularizer and uses an adaptive learning rate based on Hedge—see Algorithm 1, where $G_t = e_{i_t} g_t^\top \in \mathbb{R}^{N \times d}$ denotes the gradient matrix full of 0 except for row i_t which contains g_t. The next result bounds the regret suffered by MT-FTRL.

Theorem 1 (Cesa-Bianchi et al. (2022, Theorem 9))**.** *Let G be a clique. The regret of* MT-FTRL *with $\beta_{t-1} = \sqrt{t}$ satisfies for all $U \in \mathcal{U}$*

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1 + \sigma^2(N-1)}\sqrt{T}\,, \tag{3}$$

where $\sigma^2 = \sigma^2(U) = \frac{1}{N-1} \sum_{i=1}^N \big\|U_{i:} - \frac{1}{N}\sum_{j=1}^N U_{j:}\big\|_2^2$ is the comparator variance.[1]

Note that (3) is at most of the same order as the naive bound (2), as $\sigma^2(U) \le 8$ for all U. On the other hand, (3) gets better as σ^2 decreases, i.e., as the tasks get similar, and recovers the single-agent bound $\sqrt{T}$ when $\sigma^2 = 0$, i.e., when all tasks are identical. Finally, we stress that MT-FTRL adapts to the true comparator variance $\sigma^2(U)$ without any prior knowledge about it.

The above approach crucially relies on the fact that the communication graph G is a clique. In the next

[1] We use $g \overset{\widetilde{\mathcal{O}}}{=} f$ to denote $g = \widetilde{\mathcal{O}}(f)$, where $\widetilde{\mathcal{O}}$ hides logarithmic factors in N.

Algorithm 1 `MT-FTRL` (on linear losses)

Requires: Number of agents N, learning rates β_{t-1}

Init: $A = (1+N)I_N - \mathbf{1}_N\mathbf{1}_N^\top$, $\Xi = \{1/N, 2/N, \dots, 1\}$, $p_1^{(\xi)} = \frac{1}{N}$ for all $\xi \in \Xi$

for $t = 1, 2, \dots$ **do**

  for $\xi \in \Xi$ **do**

    // Set learning rate assuming $\sigma^2 = \xi$

    $\eta_{t-1}^{(\xi)} = \frac{N}{\beta_{t-1}}\sqrt{1+\xi(N-1)}$

    // FTRL with Mahalanobis regularizer

$$X_t^{(\xi)} = \mathop{\arg\min}_{\substack{X\in\mathcal{U}\\ \sigma^2(X)\le\xi}} \eta_{t-1}^{(\xi)} \left\langle \sum_{s\le t-1} G_s, X \right\rangle + \frac{1}{2}\|X\|_A^2$$

  // Average the experts predictions

  Predict $X_t = \sum_{\xi\in\Xi} p_t^{(\xi)} X_t^{(\xi)}$

  Incur linearized loss $\langle g_t, [X_t]_{i_t:}\rangle$ and observe g_t

  // Update p_t based on the experts losses

  for $\xi \in \Xi$ **do**

$$p_{t+1}^{(\xi)} = \frac{\exp\left(-\frac{\sqrt{\ln N}}{\beta_t}\sum_{s=1}^t \langle g_s, [X_s^{(\xi)}]_{i_s:}\rangle\right)}{\sum_k \exp\left(-\frac{\sqrt{\ln N}}{\beta_t}\sum_{s=1}^t \langle g_s, [X_s^{(k)}]_{i_s:}\rangle\right)}$$

section, we show how to use `MT-FTRL` as a building block to devise an algorithm that can operate on any communication graph. Some of our bounds depend on notable parameters of the graph G, that we recall now.

Definition 1. *Let $G = (V, E)$. A subset $S \subset V$ is*

- k-times independent *in G if the shortest path between any two vertices in S is of length $k+1$. The cardinality of the largest k-times independent set of G is denoted $\alpha_k(G)$. For $k = 1$, we use the term* independent set *and adopt the notation $\alpha(G)$.*
- dominating *in G if any vertex in $G\setminus S$ has a neighbor in S. The cardinality of the smallest dominating set of G is denoted $\gamma(G)$ and called* domination number of G.

It is well known that: $\alpha_2(G) \le \gamma(G) \le \alpha(G) \le |V|$.

3 ALGORITHM AND ANALYSIS

In this section, we introduce and analyze `MT-CO2OL`, a meta-algorithm for MultiTask COmmunication-COnstrained Online Learning. In particular, we prove two sets of regret upper bounds, depending on whether the agent activations are adversarial or stochastic. We also prove some lower bounds on the regret.

`MT-CO2OL` uses as building block any multitask algorithm able to operate when the communication graph is a clique. In the following, we generically refer to this base algorithm as `AlgoClique`. `MT-CO2OL` requires each

Algorithm 2 `MT-CO2OL`

Requires: Base algorithm `AlgoClique`, weights w_{ij}

for $t = 1, 2, \dots$ **do**

  Active agent i_t

    fetches $\left[Y_t^{(j)}\right]_{i_t:}$ from each $j \in \mathcal{N}_{i_t}$

    predicts $x_t = \sum_{j\in\mathcal{N}_{i_t}} w_{i_t j}\left[Y_t^{(j)}\right]_{i_t:}$

    pays $\ell_t(x_t)$ and observes $g_t \in \partial\ell_t(x_t)$

    sends $\left(i_t, w_{i_t j}\, g_t\right)$ to each $j \in \mathcal{N}_{i_t}$

  for $j \in \mathcal{N}_{i_t}$ **do**

    Agent j feeds the linear loss $\langle w_{i_t j}\, g_t, \cdot\rangle$ to their local instance of `AlgoClique` and obtains $Y_{t+1}^{(j)}$

agent $j \in V$ to run an instance of `AlgoClique` on a virtual clique over its neighbors $i \in \mathcal{N}_j$.[2] The instance of `AlgoClique` run by j maintains a matrix $Y_t^{(j)} \in \mathbb{R}^{N_j\times d}$, whose each row $\left[Y_t^{(j)}\right]_{i:}$ stores a model associated to task i, for any $i \in \mathcal{N}_j$. At time t, each $j \in \mathcal{N}_{i_t}$ sends to the active agent i_t (or equivalently i_t fetches from j) the model $\left[Y_t^{(j)}\right]_{i_t:}$. The prediction made by i_t at time t is the weighted average $x_t = \sum_{j\in\mathcal{N}_{i_t}} w_{i_t j}\left[Y_t^{(j)}\right]_{i_t:}$, where w_{ij} are non-negative weights satisfying $\sum_{j\in\mathcal{N}_i} w_{ij} = 1$. After predicting, i_t observes $g_t \in \partial\ell_t(x_t)$ and sends $w_{i_t j}\, g_t$ to each $j \in \mathcal{N}_{i_t}$. Agents $j \in \mathcal{N}_{i_t}$ then feed the linear loss $\langle w_{i_t j} g_t, \cdot\rangle$ to their local instance of `AlgoClique`, obtaining the updated matrix $Y_{t+1}^{(j)}$. The pseudocode of `MT-CO2OL` is summarized in Algorithm 2. See Figure 2 for an illustration of one iteration. One interpretation behind `MT-CO2OL` is as follows: by fetching the predictions maintained by their neighbors, the active agent i_t actually leverages information from the neighbors of their neighbors, i.e., from agents $k \in \mathcal{N}_j$, for $j \in \mathcal{N}_{i_t}$. Indeed, these agents have a direct influence $Y_t^{(j)}$, so in particular on $\left[Y_t^{(j)}\right]_{i_t}$. By iterating this mechanism, `MT-CO2OL` allows to propagate information along the communication graph. Note that the local instance of `AlgoClique` run by agent j does not have access to the global time t, e.g., to set the learning rate, but may only use the local time $\sum_{s\le t}\mathbb{I}\{i_s \in \mathcal{N}_j\}$.

An important consequence of the fetch and send operations in `MT-CO2OL` is that they enable running `AlgoClique` *as if each agent j were part of an isolated clique over $\mathcal{N}_j$*, which in turn makes Theorem 1 applicable. Formally, let `MT-CO2OL` be run over an arbitrary sequence $i_1, \dots, i_T$ of agent activations in a communication network $G = (V, E)$. Then, for each $j \in V$ and time t, the matrix $Y_t^{(j)}$ of models computed by `AlgoClique` run at node j is identical to the ma-

[2] Note that the alternative consisting in running `AlgoClique` for each pair $(i, j) \in E$ fails, see Appendix A.6.

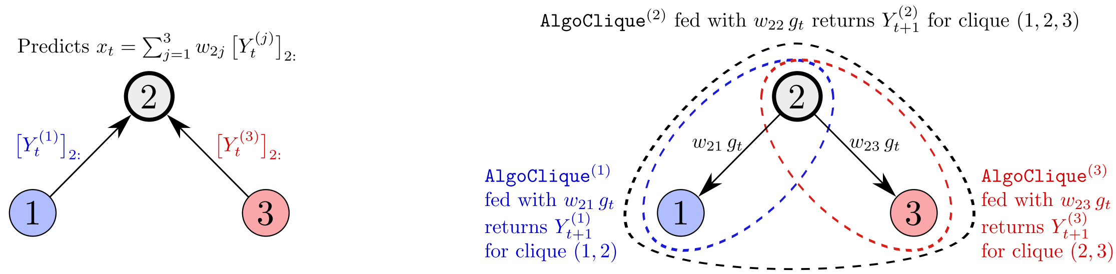


Figure 2: One Iteration of MT-CO$_2$OL. Active agent 2 fetches their neighbor's models and predicts with their weighted average x_t (left). Then 2 pays $\ell_t(x_t)$, observes $g_t \in \partial \ell_t(x_t)$, and sends back $w_{2j}\, g_t$. Finally, each local AlgoClique instance updates the models for the agents in the virtual clique centered on the agent running the instance (right).

trix computed by AlgoClique run on a clique over $\mathcal{N}_j$ with the sequence of activations restricted to $\mathcal{N}_j$. This simple yet key observation allows us to derive the following lemma, which shows that the regret of MT-CO$_2$OL can be expressed in terms of the regrets suffered by AlgoClique run on the artificial cliques $\mathcal{N}_j$, for $j \in V$.

Lemma 2. *For any $U \in \mathbb{R}^{N\times d}$, let $U^{(j)} \in \mathbb{R}^{N_j \times d}$ be the matrix whose rows contain all $U_{i:}$ for i in $\mathcal{N}_j$, sorted in ascending order of i. Furthermore, let $R_T^{\text{clique-}j}$ be the regret suffered by* AlgoClique *run by agent j on the linear losses $\langle w_{i_t j}\, g_t, \cdot\rangle$ over the rounds $t \le T$ such that $i_t \in \mathcal{N}_j$. Then, the regret of* MT-CO$_2$OL *satisfies*

$$\forall\, U \in \mathbb{R}^{N\times d}, \quad R_T(U) \le \sum_{j=1}^{N} R_T^{\text{clique-}j}\left(U^{(j)}\right).$$

Proof. Let $x_1, \ldots, x_T$ be the predictions of MT-CO$_2$OL. By convexity of the ℓ_t, we have

$$\sum_{t=1}^{T} \ell_t(x_t) - \ell_t(U_{i_t:}) \le \sum_{t=1}^{T} \langle g_t, x_t - U_{i_t:}\rangle \tag{4}$$

$$= \sum_{t=1}^{T}\sum_{i=1}^{N} \Big\langle g_t, \sum_{j\in\mathcal{N}_i} w_{ij}\left[Y_t^{(j)}\right]_{i:} - U_{i:}\Big\rangle \mathbb{I}\{i_t = i\}$$

$$= \sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i} \Big\langle g_t, w_{ij}\left(\left[Y_t^{(j)}\right]_{i:} - U_{i:}\right)\Big\rangle \mathbb{I}\{i_t = i\}$$

$$= \sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{i\in\mathcal{N}_j} \Big\langle w_{ij}\, g_t, \left[Y_t^{(j)}\right]_{i:} - U_{i:}\Big\rangle \mathbb{I}\{i_t = i\} \tag{5}$$

$$= \sum_{j=1}^{N} R_T^{\text{clique-}j}\left(U^{(j)}\right),$$

where (5) follows from Lemma 9. □

Note that we could obtain similar guarantees by using the convexity of ℓ_t and postponing the linearization step (4) to the individual clique regrets. However, this would require computing the subgradients at the partial predictions $\left[Y_t^{(j)}\right]_{i:}$. Instead, MT-CO$_2$OL only uses the subgradient g_t evaluated at the true prediction x_t. In the next section, we show how to leverage Lemma 2 to derive regret bounds for MT-CO$_2$OL.

3.1 Adversarial Activations

In this section, we assume the sequence $i_1, i_2, \ldots \in V$ of agent activations is chosen by an oblivious adversary. In the next theorem, we show that setting AlgoClique as MT-FTRL yields good regret bounds. Importantly, since the instance of MT-FTRL of agent j is run on the virtual clique $\mathcal{N}_j$, the regret bound scales with the *local* task variances $\sigma_j^2 = \frac{1}{2N_j(N_j-1)}\sum_{i,i'\in\mathcal{N}_j}\left\|U_{i:} - U_{i':}\right\|_2^2$, quantifying how similar neighboring tasks are. Let $\sigma_{\max}^2 = \max_{j\in[N]} \sigma_j^2$ and $\sigma_{\min}^2 = \min_{j\in[N]}\sigma_j^2$. We are now ready to state our result.

Theorem 3. *Let $G = (V,E)$ be any communication graph. Consider* MT-CO$_2$OL *where the base algorithm* AlgoClique *run by each agent $j \in V$ is an instance of* MT-FTRL *with parameters $N = N_j$ and $\beta_{t-1} = \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1 + \sum_{s\le t-1}\mathbb{I}\{i_s \in \mathcal{N}_j\}}$. Then, the regret of* MT-CO$_2$OL *satisfies for all $U \in \mathcal{U}$*

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sum_{j=1}^{N} \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1 + \sigma_j^2(N_j - 1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\,.$$

Setting $w_{ij} = \mathbb{I}\{j \in \mathcal{N}_i\}/N_i$ we obtain

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \min\left\{\frac{\sqrt{NN_{\max}}}{N_{\min}}\sqrt{1 + \sigma_{\max}^2(N_{\max} - 1)}\,,\ \frac{N}{N_{\min}}\sqrt{1 + \bar{\sigma}^2(N_{\max} - 1)}\right\}\sqrt{T}\,.$$

where $\bar{\sigma}^2 = (1/N)\sum_{j=1}^{N}\sigma_j^2$ is the average local variance. Finally, there exist some w_{ij} such that

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1 + \Delta^2(N_{\max} - 1)}\sqrt{\gamma(G)\, T}\,,$$

where $\Delta^2 = \sup_{(i,j)\in E}\|U_{i:} - U_{j:}\|_2^2$.

Theorem 3 provides three results with different flavors. The first bound is the most general, and holds for any choice of weights w_{ij}. It shows that the regret of MT-CO$_2$OL improves as the σ_j^2 get smaller, i.e., when neighbors in G have similar tasks. The second bound uses uniform weights that only depend on local information (the neighborhood sizes). Interestingly, the bound depends on the total horizon T instead of the individual T_i, and gets smaller as G becomes dense ($N_{\min} \gg 1$) and regular ($N_{\max}/N_{\min} \to 1$). Finally, the third bound uses knowledge of the full graph G (in particular of its smallest dominating set, which is NP-hard to compute) to set the weights. This bound can be viewed as a multitask version of the single-task regret bound $\sqrt{\gamma(G)\,T}$ (recovered for $\Delta^2 = 0$) achieved by the centralized algorithm that pre-computes this dominating set and then delegates the predictions of each node to its dominating node. These bounds also show the improvement brought by sharing models, as methods sharing only gradients suffer from a lower bound of $\sqrt{\alpha(G)\,T}$ (Cesa-Bianchi et al., 2020). We now instantiate our bound for uniform weights to some particular graphs of interest.

Corollary 4. *Under the assumptions of Theorem 3, set $w_{ij} = \mathbb{I}\{j \in \mathcal{N}_i\}/N_i$. If G is K-regular, then*

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1 + K\,\sigma_{\max}^2}\sqrt{\frac{NT}{K+1}}\,.$$

If G is a union of χ cliques, then

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{\chi + \sigma_{\max}^2 (N - \chi)}\sqrt{T}\,.$$

The regret bound for K-regular graphs improves upon the $\sqrt{\alpha(G)\,T}$ bound proven by Cesa-Bianchi et al. (2020) in the more restrictive setting of single-task problems with stochastic activations. Indeed, by Turan's theorem, see e.g., Mannor and Shamir (2011, Lemma 3), we have that $N/(K+1) \le \alpha(G)$ for K-regular graphs. This shows that the fetch step in MT-CO$_2$OL, which is missing in the learning protocol of Cesa-Bianchi et al. (2020), is key to derive improved regret guarantees. Regarding the second bound in Corollary 4, we recover the bound of MT-FTRL (Cesa-Bianchi et al., 2022) when $\chi = 1$. Notably, MT-CO$_2$OL achieves these regret bounds without requiring knowledge of the graph structure.

3.2 Stochastic Activations

In this section, we show that MT-CO$_2$OL enjoys improved regret bounds when the agent activations are stochastic rather than adversarial. Formally, let $q_1, \dots, q_N \in [0,1]^N$ such that $\sum_{i=1}^N q_i = 1$. Let $i_1, \dots, i_T$ be independent random variables denoting agent activations such that for every t we have $\mathbb{P}(i_t = i) = q_i$. We also define $Q_j = \sum_{i \in \mathcal{N}_j} q_i$. In the following theorem, we show how MT-CO$_2$OL can adapt to the stochasticity to attain better regret bounds. For simplicity, we assume the algorithm discussed below has access to the conditional probabilities q_i/Q_j. In Remark 2, we discuss a simple strategy to extend the analysis to the case where only a lower bound on $\min_i q_i$ is available.

Theorem 5. *Let $G = (V, E)$ be any communication graph. Consider MT-CO$_2$OL where the base algorithm AlgoClique run by each agent $j \in V$ is an instance of MT-FTRL with parameters $N = N_j$ and $\beta_{t-1} = \sqrt{\sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_j} w_{ij}^2}\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}$. Then, the regret of MT-CO$_2$OL satisfies for all $U \in \mathcal{U}$*

$$\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \left(\sum_{j=1}^N \sqrt{\sum_{i \in \mathcal{N}_j} q_i\, w_{ij}^2}\sqrt{1 + \sigma_j^2 (N_j - 1)}\right)\sqrt{T}\,,$$

where the expectation is taken over $i_1, \dots, i_T$. Setting $w_{ij} = q_j/Q_i$ we obtain

$$\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1 + \sigma_{\max}^2 (N_{\max} - 1)}\sqrt{\alpha(G)\,T}\,.$$

The improvement with respect to Theorem 3 is a natural consequence of the stochastic activations. Indeed, the term $\max_{i \in \mathcal{N}_j} w_{ij}$ arises because agent j receives scaled gradients from its neighbors $i \in \mathcal{N}_j$, with squared norms bounded by $\|w_{ij}\, g_t\|_2^2 \le \max_{i \in \mathcal{N}_j} w_{ij}^2$. With stochastic activations, this norm can be bounded in expectation by $\sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_j} w_{ij}^2$, resulting in the improvement seen in in the first bound of Theorem 5. In the second bound, we show that an appropriate choice of weights w_{ij} allows to recover the $\sqrt{\alpha(G)\,T}$ bound derived by Cesa-Bianchi et al. (2020) in the single-task setting ($\sigma_{\max}^2 = 0$). Importantly, the choice of w_{ij} remains local to agent j, and does not require knowledge of the full communication graph. Finally, while Corollary 4 shows regret bounds sharper than $\sqrt{\alpha(G)\,T}$ in the more general adversarial case, those bounds only apply to particular graphs. In contrast, the results of Theorem 5 are valid for any graph. We conclude this section with a remark extending the guarantees to the case where the q_j/Q_i have to be estimated.

Remark 2 (Extension to unknown q_i)**.** *Both the choice of learning rates β_{t-1} and weights w_{ij} in the second claim of Theorem 5 require that agent j can access the conditional probabilities q_i/Q_j. When the latter are unknown, they can be estimated easily through $\widehat{\pi}_{ij}(t) := \sum_{s \le t} \mathbb{I}\{i_s = i\} / \sum_{s \le t} \mathbb{I}\{i_s \in \mathcal{N}_j\}$. In Appendix A.4, we show that the 2-step strategy consisting in first computing the $\widehat{\pi}_{ij}$ and then running MT-CO$_2$OL with $\widehat{\pi}_{ij}$ instead of q_i/Q_j suffers minimal additional regret. Note that, although appealing, the idea of replacing q_i/Q_j in β_{t-1} by $\widehat{\pi}_{ij}(t)$ at each time step t would not work, as it would disrupt the monotonicity of the learning rate sequence β_{t-1}.*

3.3 Lower Bounds

We conclude the regret analysis by providing lower bounds. In particular, we show that MT-CO_2OL's regret bound for regular graphs is tight up to constant factors.

Theorem 6. *Let G be any communication graph. Then for any algorithm following Section 2's protocol:*

1. there exists a sequence of activations and gradients such that the algorithm suffers regret

$$\sup_{\substack{U\in\mathcal{U}\\ \sigma^2(U)\le\nu^2}} R_T(U) \ge \max\left\{\sqrt{1+\nu^2(N-1)}, \sqrt{\alpha_2(G)}\right\}\frac{\sqrt{T}}{3}.$$

2. for any even number K, there exists a K-regular graph and a sequence of activations and gradients such that the algorithm suffers regret

$$\sup_{U\in\mathcal{U}} R_T(U) \ge \frac{1}{5}\sqrt{1+K\,\sigma^2_{\min}}\sqrt{\frac{NT}{K}}\,.$$

The first result is the maximum of two lower bounds. The bound depending on the (upper bound of the) comparator variance ν^2 is due to Cesa-Bianchi et al. (2022), and holds even without communication constraints. The bound depending on $\alpha_2(G)$ is established in Appendix A.5, and leverages the fact that there is no information leakage if the activated nodes are always 2 edges apart. The second result investigates this idea further for K-regular graphs, and provides a matching lower bound for Corollary 4's first claim.

4 A PRIVATE VARIANT

MT-CO_2OL involves a gradient-sharing step, potentially harming privacy if losses contain sensitive information. In this section, we modify MT-CO_2OL so as to satisfy *loss-level differential privacy.* We prove that privacy only degrades the regret by a term polylogarithmic in T. Additionally, we establish privacy thresholds where sharing information becomes ineffective. Our analysis is based on the following privacy notion.

Definition 2. *We consider a randomized multi-agent algorithm $\mathcal{A}$, governing the communication between N agents. Let $m_t^{(i)}$ denote the batch of messages sent by agent i to the other agents at time t. For any $\varepsilon > 0$, $\mathcal{A}$ is loss-level ε-differentially private (or ε-DP for short) if for all $i \in [N]$ and set $\mathcal{M}$ of sequences of messages*

$$\frac{\mathbb{P}\left(m_1^{(i)},\dots,m_T^{(i)} \in \mathcal{M} \mid i_1,\ell_1,\dots,i_T,\ell_T\right)}{\mathbb{P}\left(m_1^{(i)},\dots,m_T^{(i)} \in \mathcal{M} \mid i_1,\ell'_1,\dots,i_T,\ell'_T\right)} \le e^{\varepsilon}\,, \quad (6)$$

where $(\ell'_t)_{t\le T}$ and $(\ell_t)_{t\le T}$ differ by at most one entry.

As it applies to messages between agents, our definition of DP is more general than traditional notions which focus on predictions, as commonly done in single-agent scenarios (Jain et al., 2012; Agarwal and Singh, 2017). Note that a similar definition has been employed by Bellet et al. (2018), in the batch case though.

Algorithm 3 DPMT-CO_2OL

Requires: Base algorithm DPMT-FTRL, weights w_{ij}, privacy level ε, distributions $\mathcal{D}_d$, $\mathcal{D}_1$

for $t = 1, 2, \dots$ **do**

- Active agent i_t
  - fetches $\left[Y_t^{(j)}\right]_{i_t:}, \left\{\left[X_t^{(j,\xi)}\right]_{i_t:} : \xi \in \Xi_j\right\}$ from each $j \in \mathcal{N}_{i_t}$
  - predicts $x_t = \sum_{j\in\mathcal{N}_{i_t}} w_{i_t j}\left[Y_t^{(j)}\right]_{i_t:}$
  - pays $\ell_t(x_t)$ and observes $g_t \in \partial\ell_t(x_t)$
  - updates $\gamma_t^{(i_t)} = \gamma_t^{(i_t)} + g_t$
  - computes $\widetilde{\gamma}_t^{(i_t)}$ using an instance of TreeBasedAgg with distribution $\mathcal{D}_d$
- **for** $j \in \mathcal{N}_{i_t}$ **do**
  - **for** $\xi \in \Xi_j$ **do**
    - $s_{t,i_t}^{(j,\xi)} = s_{t-1,i_t}^{(j,\xi)} + w_{i_t j}\langle [X_t^{(j,\xi)}]_{i_t:}, g_t\rangle$,
    - computes $\widetilde{s}_{t,i_t}^{(j,\xi)}$ using an instance of TreeBasedAgg set with distribution $\mathcal{D}_1$
  - sends $\left(i_t, w_{i_t j}\,\widetilde{\gamma}_t^{(i_t)}, \{\widetilde{s}_{t,i_t}^{(j,\xi)} : \xi \in \Xi_j\}\right)$ to j
  - Agent j feeds $\left(w_{i_t j}\,\widetilde{\gamma}_t^{(i_t)}, \{\widetilde{s}_{t,i_t}^{(j,\xi)} : \xi \in \Xi_j\}\right)$ to their local instance of DPMT-FTRL and obtains $\left(Y_{t+1}^{(j)}, \{X_{t+1}^{(j,\xi)},\ \forall \xi \in \Xi_j\}\right)$.

Our modification of MT-CO_2OL relies on the fact that MT-FTRL only requires information on sums of gradients and sums of inner products between predictions and gradients (for Hedge). This allows us to use aggregation trees (Dwork et al., 2010; Chan et al., 2011), enabling the DP release of cumulative vector sums so that the level of noise introduced remains logarithmic in the number of vectors. In particular, agent i_t can compute:

- Sanitized versions $\widetilde{\gamma}_t^{(i)}$ of the sums of gradients observed by the agents $\sum_{s\le t} g_s\,\mathbb{I}\{i_s = i\}$;
- Sanitized versions $\widetilde{s}_{t,i}^{(j,\xi)}$ of the sums of inner products between predictions of expert ξ of agent j and weighted gradients $\sum_{s\le t-1}\langle w_{ij} g_s, \left[X_s^{(j,\xi)}\right]_{i:}\rangle\mathbb{I}\{i_s = i\}$.

This technique was used in Guha Thakurta and Smith (2013); Agarwal and Singh (2017) for DP single-agent online learning. As in Agarwal and Singh (2017), we use TreeBasedAgg, which tweaks the original tree-aggregation algorithm to obtain identically distributed noise across rounds.

The resulting algorithm, called DPMT-CO_2OL, is described in Algorithm 3. It invokes the ε-DP version of

MT-FTRL, summarized in the supplementary material (Algorithm 4). We now prove that $\texttt{DPMT-CO}_2\texttt{OL}$ is ε-DP.

Theorem 7. *Let G be any graph, and for any $\varepsilon > 0$, let $\varepsilon' = \varepsilon/(6N_{\max}^2)$. Assume that $\texttt{DPMT-CO}_2\texttt{OL}$ is run with $\mathcal{D}_d$ set to a d-dimensional Laplacian distribution with parameter $\frac{\sqrt{d}\ln T}{\varepsilon'}$, and $\mathcal{D}_1$ set to a 1-dimensional Laplacian distribution with parameter $\frac{\ln T}{\varepsilon'}$. Then $\texttt{DPMT-CO}_2\texttt{OL}$ is ε-DP.*

Next, we state the regret bound for $\texttt{DPMT-CO}_2\texttt{OL}$ with adversarial agent activations.

Theorem 8. *Let $G = (V, E)$ be any communication graph, Consider $\texttt{DPMT-CO}_2\texttt{OL}$ where the base algorithm run by each agent $j \in V$ is an instance of DPMT-FTRL with parameters $N = N_j$ and $\beta_{t-1} = \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1 + \sum_{s\le t-1}\mathbb{I}\{i_s \in \mathcal{N}_j\}}$. Then the regret of $\texttt{DPMT-CO}_2\texttt{OL}$ with $\mathcal{D}_d$ and $\mathcal{D}_1$ set as in Theorem 7 satisfies*

$$\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \sum_{j=1}^{N} \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i} + \frac{dN_{\max}^4}{\varepsilon} N \ln^2 T\,, \tag{7}$$

where randomness is due to sanitization.

Theorem 8 shows that the additional regret due to the privatization of our algorithm (second line in bound (7)) has a very mild dependence on T. In particular, when G is K-regular, it becomes negligible as soon as $T \ge N(d^2K^9/\varepsilon^2)$. On the other side, the bound indicates that running $\texttt{DPMT-CO}_2\texttt{OL}$ is worse than running N instances of FTRL without communication (which is DP by design) when $\varepsilon \le dN_{\max}^4\sqrt{N/T}$. In short, if agents are too private, listening to neighbors introduces more noise than valuable information. Note that the cut-off value for ε below which sharing information is detrimental vanishes as T goes to $+\infty$. Hence, for any privacy level ε, $\texttt{DPMT-CO}_2\texttt{OL}$ is asymptotically (in T) preferable to independent instances of FTRL. Note finally that the cost of privacy remains the same for stochastic activations, and that in the single-agent case we do recover the result of Agarwal and Singh (2017).

Another possibility to maintain DP is to implement $\texttt{DPMT-CO}_2\texttt{OL}$ by adding Laplacian noise of parameter $1/\varepsilon$ to each individual gradient. It can be shown that this method yields a total privacy cost of order at least $d\left(\sum_{j=1}^N \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\right)/\varepsilon$. In particular, when G is K-regular, the cut-off value for ε becomes $d/\sqrt{K}$, which does not vanish as $T \to \infty$. Hence, this way of privatizing $\texttt{MT-CO}_2\texttt{OL}$ is too coarse to maintain better performances over non-communicating methods for all privacy levels ε.

Interestingly, the proof of Theorem 8 also shows that knowing the task variances σ_j^2 reduces the cost of privacy, because the instances of MT-FTRL do not need to use experts to adapt to σ_j^2. In particular, the active agent can avoid sharing $\widetilde{s}_{t,i_t}^{(j,\xi)}$, which leads to a significant reduction in the information leakage. Then, the cost of privacy becomes of order $\left(\sum_{j=1}^N \sqrt{N_j\left(1+\sigma_j^2(N_j-1)\right)}\right) d\ln^2 T/\varepsilon$.

5 EXPERIMENTS

In this section, we describe experiments supporting our theoretical claims. In our experiments, the communication graphs G are generated using the Erdős–Rényi's model: given $N = 30$ vertices, the edge between two vertices is included in G with probability $p = 0.9$. Our results are averaged over 48 independent draws of such graphs. The losses are of the form $\ell_t(x) = \frac{1}{2}\|x - z_t\|_2^2$, where the z_t are realizations of a multivariate Gaussian with covariance matrix $10^{-4}\, I_d$ and expectation $U_{i_t:}$. Here, i_t denotes the active task, and $U_{i_t:}$ the corresponding task vector. The matrix U is drawn according to a centered Gaussian distribution with covariance matrix $(I_N + \lambda L_G)^{-1}$, where L_G is the Laplacian of the graph G. The local task variance is then controlled via the parameter λ, which in our experiments varies in the set $\{10^{10}, 10, 9, 8, 7, 6, 5, 4, 3, 2\}$.[3] The average local task standard deviation $\bar{\sigma}$ resulting from these choices of λ are the x-coordinates of points in Figure 3. Activations are stochastically generated with $q_i = 1/N$.

For practical reasons, the version of $\texttt{MT-CO}_2\texttt{OL}$ used here slightly differs from Algorithm 2, as our adaptive part is based on the Krichevsky-Trofimov algorithm (see Algorithm 5 for details) to account for task variances and domain diameter. This variant has regret guarantees similar to the original $\texttt{MT-CO}_2\texttt{OL}$ (see Appendix C) but is simpler to run in practice. It is however nontrivial to make it ε-DP, which is the reason why in the theoretical analysis we privileged Hedge over Krichevsky-Trofimov. We consider two baselines: the first one, i-FTRL, consists in running instances of FTRL on agents that do not communicate with each other. The second method, ST-FTRL, is the multi-agent single-task algorithm with communication graph G introduced in (Cesa-Bianchi et al., 2020). Note that this algorithm requires oracle access to the loss function, instead of just access to the gradient computed at the current prediction. This represents a significant advantage. In particular, the regret guarantees for this algorithm only hold given access to the loss function oracle. To ensure fair comparisons, we extend adaptivity to the diameter to both i-FTRL and ST-FTRL using the Krichevsky-Trofimov method.

[3] We mimicked $\lambda = \infty$, i.e., no task variance, by $\lambda = 10^{10}$.

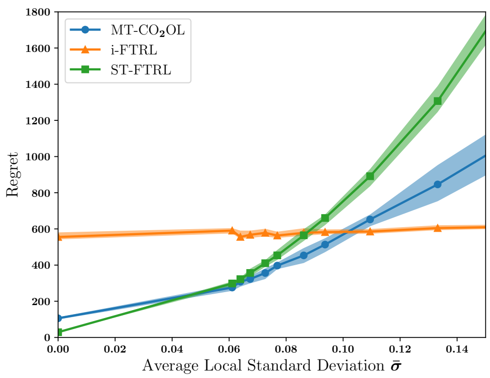


Figure 3: Multitask regret at horizon $T = 150\,000$.

Figure 3 shows the (multitask) regret of these algorithms. As expected, the performance of `i-FTRL` is independent from $\bar{\sigma}$ and dominates when the latter is large (i.e., tasks are very different). On the other hand, `ST-FTRL` wins when $\bar{\sigma}$ is close to zero (i.e., tasks are very similar). Finally, `MT-CO`$_2$`OL` outperforms the baselines for intermediate values of $\bar{\sigma}$ (see also Figure 1 that plots the regret against time for $\bar{\sigma} = 0.08$).

6 CONCLUSION

We introduced and analyze `MT-CO`$_2$`OL`, an algorithm tackling multitask online learning on arbitrary communication networks. Interesting generalizations of this work may focus on the cases where: (1) the communication network changes over time, (2) privacy levels are user-specific, (3) agents have further constraints (e.g., size-wise or frequency-wise) on the messages they send.

Acknowledgements

The authors acknowledge the financial support from the MUR PRIN grant 2022EKNE5K (Learning in Markets and Society), the NextGenerationEU program within the PNRR-PE-AI scheme (project FAIR), the EU Horizon CL4-2021-HUMAN-01 research and innovation action under grant agreement 101070617 (project ELSA). The authors thank anonymous reviewer 2 for insightful comments that helped improve the correctness of the paper.

References

Agarwal, N. and Singh, K. (2017). The price of differential privacy for online learning. In *International Conference on Machine Learning*, pages 32–40. PMLR.

Awerbuch, B. and Kleinberg, R. (2008). Competitive collaborative learning. *Journal of Computer and System Sciences*, 74(8):1271–1288.

Bar-On, Y. and Mansour, Y. (2019). Individual regret in cooperative nonstochastic multi-armed bandits. *Advances in Neural Information Processing Systems*, 32.

Bellet, A., Guerraoui, R., Taziki, M., and Tommasi, M. (2018). Personalized and private peer-to-peer machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 473–481. PMLR.

Cavallanti, G., Cesa-Bianchi, N., and Gentile, C. (2010). Linear algorithms for online multitask classification. *The Journal of Machine Learning Research*, 11:2901–2934.

Cesa-Bianchi, N., Cesari, T., and Monteleoni, C. (2020). Cooperative online learning: Keeping your neighbors updated. In *Algorithmic learning theory*, pages 234–250. PMLR.

Cesa-Bianchi, N., Gentile, C., Mansour, Y., et al. (2019). Delay and cooperation in nonstochastic bandits. *Journal of Machine Learning Research*, 20(17):1–38.

Cesa-Bianchi, N., Laforgue, P., Paudice, A., et al. (2022). Multitask online mirror descent. *Transactions on Machine Learning Research*.

Chan, T.-H. H., Shi, E., and Song, D. (2011). Private and continual release of statistics. *ACM Transactions on Information and System Security (TISSEC)*, 14(3):1–24.

Della Vecchia, R. and Cesari, T. (2021). An efficient algorithm for cooperative semi-bandits. In *Algorithmic Learning Theory*, pages 529–552. PMLR.

Dwork, C., Naor, M., Pitassi, T., and Rothblum, G. N. (2010). Differential privacy under continual observation. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 715–724.

Griggs, J. R. (1983). Lower bounds on the independence number in terms of the degrees. *Journal of Combinatorial Theory, Series B*, 34(1):22–39.

Guha Thakurta, A. and Smith, A. (2013). (nearly) optimal algorithms for private online learning in full-information and bandit settings. *NeurIPS*, 26.

Herbster, M., Pasteris, S., Vitale, F., and Pontil, M. (2021). A gang of adversarial bandits. *Advances in Neural Information Processing Systems*, 34:2265–2279.

Hosseini, S., Chapman, A., and Mesbahi, M. (2013). Online distributed optimization via dual averaging. In *52nd IEEE Conference on Decision and Control*, pages 1484–1489. IEEE.

Hsieh, Y.-G., Iutzeler, F., Malick, J., and Mertikopoulos, P. (2022). Multi-agent online optimization with delays: Asynchronicity, adaptivity, and optimism. *The Journal of Machine Learning Research*, 23(1):3377–3425.

Ito, S., Hatano, D., Sumita, H., Takemura, K., Fukunaga, T., Kakimura, N., and Kawarabayashi, K.-I. (2020). Delay and cooperation in nonstochastic linear bandits. *Advances in Neural Information Processing Systems*, 33:4872–4883.

Jain, P., Kothari, P., and Thakurta, A. (2012). Differentially private online learning. In *Conference on Learning Theory*, pages 24–1.

Jiang, J., Zhang, W., Gu, J., and Zhu, W. (2021). Asynchronous decentralized online learning. *Advances in Neural Information Processing Systems*, 34:20185–20196.

Li, R., Ma, F., Jiang, W., and Gao, J. (2019). Online federated multitask learning. In *Proceedings of the 7th IEEE International Conference on Big Data*, pages 215–220.

Mannor, S. and Shamir, O. (2011). From bandits to experts: On the value of side-observations. *Advances in Neural Information Processing Systems*, 24.

Mitzenmacher, M. and Upfal, E. (2005). *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press.

Murugesan, K., Liu, H., Carbonell, J., and Yang, Y. (2016). Adaptive smoothed online multi-task learning. In *Proceedings of the 29th Annual Conference on Advances in Neural Information Processing Systems*, pages 4296–4304.

Orabona, F. (2019). A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*.

Saha, A., Rai, P., Daumé, H., Venkatasubramanian, S., et al. (2011). Online learning of multiple tasks and their relationships. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, pages 643–651.

Sinha, A. and Vaze, R. (2023). Playing in the dark: No-regret learning with adversarial constraints. *arXiv preprint arXiv:2310.18955*.

Smith, V., Chiang, C.-K., Sanjabi, M., and Talwalkar, A. S. (2017). Federated multi-task learning. *NeurIPS*, 30.

Van der Hoeven, D., Hadiji, H., and van Erven, T. (2022). Distributed online learning for joint regret with communication constraints. In *International Conference on Algorithmic Learning Theory*, pages 1003–1042. PMLR.

Yan, F., Sundaram, S., Vishwanathan, S., and Qi, Y. (2012). Distributed autonomous online learning: Regrets and intrinsic privacy-preserving properties. *IEEE Transactions on Knowledge and Data Engineering*, 25(11):2483–2493.

Yi, J. and Vojnović, M. (2023). Doubly adversarial federated bandits. *arXiv preprint arXiv:2301.09223*.

Zhang, C., Zhao, P., Hao, S., Soh, Y. C., Lee, B. S., Miao, C., and Hoi, S. C. (2018). Distributed multi-task classification: a decentralized online learning approach. *Machine Learning*, 107(4):727–747.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [No]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [No]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Not Applicable]

   The code will be released upon acceptance.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:

(a) Citations of the creator If your work uses existing assets. [Not Applicable]

(b) The license information of the assets, if applicable. [Not Applicable]

(c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]

(d) Information about consent from data providers/curators. [Not Applicable]

(e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

(a) The full text of instructions given to participants and screenshots. [Not Applicable]

(b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]

(c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

A Technical Proofs (Results from Section 3)

In this section, we gather the technical proofs of the results exposed in Section 3. We start by stating a lemma which allows to invert indices when summation is made over the edges of G. This result is in particular used to prove Lemma 2. We then prove all results stated in Section 3. Finally, we analyze an unsuccessful approach that runs $|E|$ instances of `MT-FTRL` that maintain predictions for all pair of agents (i,j) such that $(i,j) \in E$.

Lemma 9. *Let $F \in \mathbb{R}^{N\times N}$ be any matrix (i.e., not necessarily symmetric). Then we have*

$$\sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} F_{ij} = \sum_{j=1}^{N} \sum_{i \in \mathcal{N}_j} F_{ij}\,.$$

Proof. We have

$$\sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} F_{ij} = \sum_{i=1}^{N}\sum_{j=1}^{N} F_{ij}\,\mathbb{I}\left\{j \in \mathcal{N}_i\right\} = \sum_{i=1}^{N}\sum_{j=1}^{N} F_{ij}\,\mathbb{I}\left\{i \in \mathcal{N}_j\right\} = \sum_{j=1}^{N} \sum_{i \in \mathcal{N}_j} F_{ij}\,.$$

□

A.1 Proof of Theorem 3

Theorem 3. *Let $G = (V,E)$ be any communication graph. Consider `MT-CO2OL` where the base algorithm `AlgoClique` run by each agent $j \in V$ is an instance of `MT-FTRL` with parameters $N = N_j$ and $\beta_{t-1} = \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sum_{s\le t-1}\mathbb{I}\left\{i_s \in \mathcal{N}_j\right\}}$. Then, the regret of `MT-CO2OL` satisfies for all $U \in \mathcal{U}$*

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sum_{j=1}^{N} \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\,.$$

Setting $w_{ij} = \mathbb{I}\left\{j \in \mathcal{N}_i\right\}/N_i$ we obtain

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \min\left\{\frac{\sqrt{NN_{\max}}}{N_{\min}}\sqrt{1+\sigma_{\max}^2(N_{\max}-1)}\,,\right.$$
$$\left.\frac{N}{N_{\min}}\sqrt{1+\bar{\sigma}^2(N_{\max}-1)}\right\}\sqrt{T}\,.$$

where $\bar{\sigma}^2 = (1/N)\sum_{j=1}^{N}\sigma_j^2$ is the average local variance. Finally, there exist some w_{ij} such that

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1+\Delta^2(N_{\max}-1)}\sqrt{\gamma(G)\,T}\,,$$

where $\Delta^2 = \sup_{(i,j)\in E}\|U_{i:}-U_{j:}\|_2^2$.

Proof. By Lemma 2, we have

$$R_T(U) \le \sum_{j=1}^{N} R_T^{\text{clique-}j}\left(U^{(j)}\right)\,,$$

where $R_T^{\text{clique-}j}$ the regret suffered by `MT-FTRL` on the linear losses $\langle w_{i_t j}\, g_t, \cdot\rangle$ over the rounds $t \le T$ such that $i_t \in \mathcal{N}_j$. We now upper bound each of these terms individually. Let $j \in [N]$, using the notation in Algorithm 1 we have

$$\begin{aligned}
R_T^{\text{clique-}j}\left(U^{(j)}\right) &= \sum_{t=1}^{T}\sum_{i\in\mathcal{N}_j}\left\langle w_{ij}\, g_t, \left[Y_t^{(j)}\right]_{i:} - U_{i:}^{(j)}\right\rangle \mathbb{I}\left\{i_t = i\right\} \\
&= \sum_{t:\, i_t\in\mathcal{N}_j}\left\langle w_{i_t j}\, g_t, \left[Y_t^{(j)}\right]_{i_t:} - U_{i_t:}^{(j)}\right\rangle
\end{aligned}$$

$$= \sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t\right]_{i_t:} - U^{(j)}_{i_t:} \right\rangle \tag{8}$$

$$\leq \underbrace{\sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t\right]_{i_t:} - \left[X_t^{(\xi^*)}\right]_{i_t:} \right\rangle}_{\text{Regret of } \texttt{Hedge}} + \underbrace{\sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t^{(\xi^*)}\right]_{i_t:} - U^{(j)}_{i_t:} \right\rangle}_{\text{Regret with choice } \xi^*}, \tag{9}$$

where $\xi^* = \arg\min_{\xi \in \Xi_j} \sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t^{(\xi)}\right]_{i_t:} \right\rangle$. We start by upper bounding the regret due to `Hedge`. Let $\text{loss}_t \in \mathbb{R}^{N_j}$ be the vector storing the $w_{i_t j} \langle g_t, \left[X_t^{(\xi)}\right]_{i_t:} \rangle$ for $\xi \in \Xi_j$, and $e^* \in \mathbb{R}^{N_j}$ the one-hot vector with an entry of 1 at expert ξ^*. By the analysis of `Hedge` with regularizers $\psi_t(\boldsymbol{p}) = \frac{\beta\sqrt{1+\sum_{s\leq t-1}\mathbb{I}\{i_s\in\mathcal{N}_j\}}}{\sqrt{\ln N_j}} \sum_{k=1}^{N_j} p_k \ln(p_k)$, see e.g., Orabona (2019, Section 7.5), we have

$$\begin{aligned}
\sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t\right]_{i_t:} - \left[X_t^{(\xi^*)}\right]_{i_t:} \right\rangle &= \sum_{t:\, i_t \in \mathcal{N}_j} \left\langle \text{loss}_t, \boldsymbol{p}_t - e^* \right\rangle \\
&\leq \beta \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{t:\, i_t \in \mathcal{N}_j} \frac{\|\text{loss}_t\|_\infty^2}{\sqrt{1 + \sum_{s \leq t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \\
&\leq \beta \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{t:\, i_t \in \mathcal{N}_j} \frac{w_{i_t j}^2}{\sqrt{1 + \sum_{s \leq t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \\
&\leq \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} \left(\beta + \frac{1}{\beta} \max_{i \in \mathcal{N}_j} w_{ij}^2\right) \\
&= 2 \max_{i \in \mathcal{N}_j} w_{ij} \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i}
\end{aligned} \tag{10}$$

if we choose $\beta = \max_{i \in \mathcal{N}_j} w_{ij}$. We now turn to the second regret. Assume first that $\sigma_j^2 \leq 1$. Then, the regret with choice ξ^* is in particular better than the regret with choice $\bar{\xi} \in \Xi_j$ such that $\bar{\xi} - \frac{1}{N_j} \leq \sigma_j^2 \leq \bar{\xi}$. Recall that the sequence $X_t^{(\bar{\xi})}$ is generated by `FTRL` with the sequence of regularizers $\frac{1}{2}\|\cdot\|_{A_j}^2 / \eta_{t-1}^{(\bar{\xi})}$. By the analysis of `FTRL`, see e.g., Orabona (2019, Corollary 7.9), we have

$$\begin{aligned}
&\sum_{t:\, i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \left[X_t^{(\bar{\xi})}\right]_{i_t:} - U^{(j)}_{i_t:} \right\rangle \\
&\qquad \leq \frac{\left\|U^{(j)}\right\|_{A_j}^2}{2\,\eta^{(\bar{\xi})}_{\sum_{i \in \mathcal{N}_j} T_i - 1}} + \frac{1}{2} \sum_{t:\, i_t \in \mathcal{N}_j} \eta_{t-1}^{(\bar{\xi})} \left\|w_{i_t j}\, G_t\right\|_{A_j^{-1}}^2 \\
&\qquad \leq \frac{1 + \sigma_j^2 (N_j - 1)}{2\sqrt{1 + \bar{\xi}(N_j - 1)}} \max_{i \in \mathcal{N}_j} w_{ij} \sqrt{\sum_{i \in \mathcal{N}_j} T_i} + \frac{N_j \sqrt{1 + \bar{\xi}(N_j - 1)}}{2 \max_{i \in \mathcal{N}_j} w_{ij}} \sum_{t:\, i_t \in \mathcal{N}_j} \frac{w_{i_t j}^2 \left[A_j^{-1}\right]_{i_t i_t} \|g_t\|_2^2}{\sqrt{1 + \sum_{s \leq t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \\
&\qquad \leq \frac{5}{2} \max_{i \in \mathcal{N}_j} w_{ij} \sqrt{1 + \bar{\xi}(N_j - 1)} \sqrt{\sum_{i \in \mathcal{N}_j} T_i} \\
&\qquad \leq \frac{5}{2} \max_{i \in \mathcal{N}_j} w_{ij} \sqrt{1 + \left(\sigma_j^2 + \frac{1}{N_j}\right)(N_j - 1)} \sqrt{\sum_{i \in \mathcal{N}_j} T_i} \\
&\qquad \leq 4 \max_{i \in \mathcal{N}_j} w_{ij} \sqrt{1 + \sigma_j^2 (N_j - 1)} \sqrt{\sum_{i \in \mathcal{N}_j} T_i}\,,
\end{aligned} \tag{11}$$

where we used $[A_j{}^{-1}]_{ii} = \frac{2}{N_j + 1}$ (see for example computations in Appendix A.2 of Cesa-Bianchi et al. (2022)) for the third inequality. Assume now that $\sigma_j^2 \geq 1$. Then, the regret with choice ξ^* is in particular better than the regret with choice 1. The latter corresponds to independent learning (Cesa-Bianchi et al., 2022) and an analysis similar to the one above shows that its regret is bounded by

$$\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{N_j\sum_{i\in\mathcal{N}_j}T_i} \le \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j}T_i}\,. \tag{12}$$

Substituting (10) and (11) or (12) (depending on the value of σ_j^2) into (9), we obtain

$$R_T(U) \le 6\sum_{j=1}^N \max_{i\in\mathcal{N}_j} w_{ij}\left(\sqrt{1+\sigma_j^2(N_j-1)}+\ln N_j\right)\sqrt{\sum_{i\in\mathcal{N}_j}T_i}\,.$$

For the second claim, substituting $w_{ij}=\mathbb{I}\{j\in\mathcal{N}_i\}/N_i$ yields

$$\begin{aligned}
\sum_{j=1}^N \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j}T_i} &= \sum_{j=1}^N \frac{\sqrt{1+\sigma_j^2(N_j-1)}}{\min_{i\in\mathcal{N}_j}N_i}\sqrt{\sum_{i\in\mathcal{N}_j}T_i} & (13)\\
&\le \frac{\sqrt{1+\sigma_{\max}^2(N_{\max}-1)}}{N_{\min}}\sum_{j=1}^N\sqrt{\sum_{i\in\mathcal{N}_j}T_i}\\
&\le \frac{\sqrt{1+\sigma_{\max}^2(N_{\max}-1)}}{N_{\min}}\sqrt{N\sum_{j=1}^N\sum_{i\in\mathcal{N}_j}T_i} & (14)\\
&= \frac{\sqrt{1+\sigma_{\max}^2(N_{\max}-1)}}{N_{\min}}\sqrt{N\sum_{i=1}^N\sum_{j\in\mathcal{N}_i}T_i}\\
&\le \frac{\sqrt{NN_{\max}}}{N_{\min}}\sqrt{1+\sigma_{\max}^2(N_{\max}-1)}\sqrt{T}\,, & (15)
\end{aligned}$$

where (14) comes from Jensen's inequality, and the following equality from Lemma 9. Starting from (13) again, we also have

$$\sum_{j=1}^N \frac{\sqrt{1+\sigma_j^2(N_j-1)}}{\min_{i\in\mathcal{N}_j}N_i}\sqrt{\sum_{i\in\mathcal{N}_j}T_i} \le \sum_{j=1}^N\sqrt{1+\sigma_j^2(N_{\max}-1)}\frac{\sqrt{T}}{N_{\min}} \le \frac{N}{N_{\min}}\sqrt{1+\bar{\sigma}^2(N_{\max}-1)}\sqrt{T}\,, \tag{16}$$

where we have used Jensen's inequality, and $\bar{\sigma}^2=(1/N)\sum_{j=1}^N\sigma_j^2$ is the average local variance. Combining (15) and (16) gives the second claim of Theorem 3.

To prove the third claim, let $S_\gamma(G)$ be a smallest dominant set of G. For each $i\in[N]$, let $j(i)$ be the node in $\mathcal{N}_i$ that belongs to $S_\gamma(G)$ (if there are several, take the one with the smallest index). We set $w_{ij}=\delta_{jj(i)}$, i.e., agent i completely delegates its prediction to its neighbour in the dominant set. Substituting $w_{ij}=\delta_{jj(i)}$ into (8), we obtain

$$R_T^{\text{clique-}j}(U^{(j)}) = \sum_{t\colon i_t\in\mathcal{V}_j}\left\langle g_t, [X_t]_{i_t:}-U_{i_t:}^{(j)}\right\rangle,$$

where $\mathcal{V}_j=\{i\in[N]\colon j(i)=j\}$ is the set of agents with j as referent node. Unrolling the proof of the first claim (substituting $\mathcal{N}_j$ by $\mathcal{V}_j$ and w_{ij} by $\delta_{jj(i)}$), we get

$$R_T(U) \overset{\widetilde{\mathcal{O}}}{\le} \sum_{j=1}^N\max_{i\in\mathcal{V}_j}\delta_{jj(i)}\sqrt{1+\widetilde{\sigma}_j^2(|\mathcal{V}_j|-1)}\sqrt{\sum_{i\in\mathcal{V}_j}T_i} = \sum_{j\in S_\gamma(G)}\sqrt{1+\widetilde{\sigma}_j^2(|\mathcal{V}_j|-1)}\sqrt{\sum_{i\colon j(i)=i}T_i}\,,$$

where $\widetilde{\sigma}_j^2$ is the local variance at j computed among its neighbors in $\mathcal{V}_j$ (rather than $\mathcal{N}_j$). Note that $\widetilde{\sigma}_j^2$ might be larger than $\sigma_{\max}^2$, but is always bounded by $\Delta^2 \coloneqq \sup_{(i,j)\in E}\|U_{i:}-U_{j:}\|_2^2$. The proof is concluded by observing that Cauchy-Schwarz inequality gives

$$\sum_{j\in S_\gamma(G)}\sqrt{1+\widetilde{\sigma}_j^2(|\mathcal{V}_j|-1)}\sqrt{\sum_{i\colon j(i)=i}T_i} \le \sqrt{1+\Delta^2(N_{\max}-1)}\sum_{j\in S_\gamma(G)}\sqrt{\sum_{i\colon j(i)=j}T_i}$$

$$
\begin{aligned}
&\le \sqrt{1+\Delta^2(N_{\max}-1)}\sqrt{|S_\gamma(G)| \sum_{j \in S_\gamma(G)} \sum_{i\,:\, j(i)=j} T_i} \\
&= \sqrt{1+\Delta^2(N_{\max}-1)}\,\sqrt{\gamma(G)\,T}\,.
\end{aligned}
$$

□

A.2 Proof of Corollary 4

Corollary 4. *Under the assumptions of Theorem 3, set $w_{ij} = \mathbb{I}\{j \in \mathcal{N}_i\}/N_i$. If G is K-regular, then*

$$
R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1+K\,\sigma_{\max}^2}\sqrt{\frac{NT}{K+1}}\,.
$$

If G is a union of χ cliques, then

$$
R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{\chi+\sigma_{\max}^2(N-\chi)}\sqrt{T}\,.
$$

Proof. The first claim of Corollary 4 is proved by using the last claim of Theorem 3, and recalling that for a K-regular graph we have $N_{\max} = N_{\min} = K+1$. Consider now a collection of χ cliques, $C_1, \dots, C_\chi$. Starting from the right-hand side of Equation (13), we have

$$
\begin{aligned}
\sum_{j=1}^{N} \frac{\sqrt{1+\sigma_j^2(N_j-1)}}{\max_{i\in\mathcal{N}_j} N_i}\sqrt{\sum_{i\in\mathcal{N}_j} T_i} &= \sum_{k=1}^{\chi}\sum_{j\in C_k} \frac{\sqrt{1+\sigma_j^2(|C_k|-1)}}{|C_k|}\sqrt{\sum_{i\in C_k} T_i} \\
&= \sum_{k=1}^{\chi} \sqrt{1+\sigma_{\max}^2(|C_k|-1)}\sqrt{\sum_{i\in C_k} T_i} \\
&\le \sqrt{\sum_{k=1}^{\chi}\big(1+\sigma_{\max}^2(|C_k|-1)\big)}\,\sqrt{\sum_{k=1}^{\chi}\sum_{i\in C_k} T_i} \\
&= \sqrt{\chi+\sigma_{\max}^2(N-\chi)}\,\sqrt{T}\,.
\end{aligned}
$$

□

A.3 Proof of Theorem 5

Theorem 5. *Let $G = (V,E)$ be any communication graph. Consider* `MT-CO2OL` *where the base algorithm* `AlgoClique` *run by each agent $j \in V$ is an instance of* `MT-FTRL` *with parameters $N = N_j$ and $\beta_{t-1} = \sqrt{\sum_{i\in\mathcal{N}_j} \frac{q_i}{Q_j} w_{ij}^2}\sqrt{1+\sum_{s\le t-1}\mathbb{I}\{i_s \in \mathcal{N}_j\}}$. Then, the regret of* `MT-CO2OL` *satisfies for all $U \in \mathcal{U}$*

$$
\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \left(\sum_{j=1}^{N}\sqrt{\sum_{i\in\mathcal{N}_j} q_i\, w_{ij}^2}\sqrt{1+\sigma_j^2(N_j-1)}\right)\sqrt{T}\,,
$$

where the expectation is taken over $i_1, \dots, i_T$. Setting $w_{ij} = q_j/Q_i$ we obtain

$$
\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \sqrt{1+\sigma_{\max}^2(N_{\max}-1)}\sqrt{\alpha(G)\,T}\,.
$$

Proof. The proof follows that of Theorem 3 until the decomposition of Equation (9). The analysis of `Hedge` with regularizers $\psi_t(\boldsymbol{p}) = \frac{\beta\sqrt{1+\sum_{s\le t-1}\mathbb{I}\{i_s\in\mathcal{N}_j\}}}{\sqrt{\ln N_j}} \sum_{k=1}^{N_j} p_k \ln(p_k)$ then gives

$$
\sum_{t\,:\, i_t\in\mathcal{N}_j} \Big\langle w_{i_t j}\, g_t, \big[X_t\big]_{i_t:} - \big[X_t^{(\xi^*)}\big]_{i_t:}\Big\rangle = \sum_{t\,:\, i_t\in\mathcal{N}_j} \big\langle \mathrm{loss}_t, \boldsymbol{p}_t - e^* \big\rangle
$$

$$\le \beta \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{t\colon i_t \in \mathcal{N}_j} \frac{\|\mathrm{loss}_t\|_\infty^2}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}}$$

$$\le \beta \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{t=1}^{T} \frac{w_{i_t j}^2 \, \mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}}\,. \tag{17}$$

Taking expectation on both sides, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sum_{t\colon i_t \in \mathcal{N}_j} \left\langle w_{i_t j}\, g_t, \big[X_t\big]_{i_t:} - \big[X_t^{(\xi^*)}\big]_{i_t:} \right\rangle \right] \\
&\quad \le \mathbb{E}\left[\beta \sqrt{\ln N_j \sum_{i \in \mathcal{N}_j} T_i} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{t=1}^{T} \frac{w_{i_t j}^2 \, \mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \right] \\
&\quad \le \beta \sqrt{\ln N_j \; \mathbb{E}\left[\sum_{i \in \mathcal{N}_j} T_i \right]} + \frac{\sqrt{\ln N_j}}{2\beta} \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[\frac{w_{i_t j}^2 \, \mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \,\middle|\, i_1, \dots, i_{t-1}, i_t \in \mathcal{N}_j \right] \right] \\
&\quad \le \beta \sqrt{Q_j T \, \ln N_j} + \frac{\sqrt{\ln N_j}}{2\beta} \sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_j} \omega_{ij}^2 \; \mathbb{E}\left[\sum_{t=1}^{T} \frac{\mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \right] \\
&\quad \le \beta \sqrt{Q_j T \, \ln N_j} + \frac{\sqrt{\ln N_j}}{\beta} \sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_j} \omega_{ij}^2 \; \mathbb{E}\left[\sqrt{\sum_{i \in \mathcal{N}_j} T_i} \right] \\
&\quad \le 2 \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} \sqrt{T \, \ln N_j}
\end{aligned}$$

if we choose $\beta = \sqrt{\sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_j} w_{ij}^2}$. Note that the same analysis can be applied to the second regret in the decomposition of Equation (9). Overall, we obtain that for all $U \in \mathcal{U}$ we have

$$\mathbb{E}[R_T(U)] \le 6 \left(\sum_{j=1}^{N} \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} \sqrt{1 + \sigma_j^2 (N_j - 1)} \right) \sqrt{T \ln N}\,.$$

To prove the second claim, substitute $w_{ij} = q_j / Q_i$ in the above equation and observe that

$$\begin{aligned}
\sum_{j=1}^{N} \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} &= \sum_{j=1}^{N} \sqrt{\sum_{i \in \mathcal{N}_j} \frac{q_i q_j^2}{Q_i^2}} = \sum_{j=1}^{N} q_j \sqrt{\sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_i^2}} \\
&\le \sqrt{\sum_{j=1}^{N} q_j \sum_{i \in \mathcal{N}_j} \frac{q_i}{Q_i^2}} = \sqrt{\sum_{i=1}^{N} \sum_{j \in \mathcal{N}_i} q_j \frac{q_i}{Q_i^2}} = \sqrt{\sum_{i=1}^{N} \frac{q_i}{Q_i}} \le \sqrt{\alpha(G)}\,,
\end{aligned}$$

where the first inequality comes from Jensen's inequality, and the second one from a known combinatorial result, see e.g., Griggs (1983) or Cesa-Bianchi et al. (2020, Lemma 3). □

A.4 Extension of Theorem 5 to unknown q_i

Theorem 10. *Let G be any graph, and assume that the agent activations are stochastic. Consider the following strategy, run independently by each agent j for its artificial clique, and based on its local time. First, predict $Y_t^{(j)} = 0$ until the local time reaches $\tau := \left\lceil \frac{4}{q_{\min}} \ln\left(2N^2 T\right) \right\rceil$, where $q_{\min} = \min_{i \in [N]} q_i$. Then, predict $Y_t^{(j)}$ using* `MT-FTRL`*, run with $N = N_j$ and $\beta_{t-1} = \sqrt{\sum_{i \in \mathcal{N}_j} \widehat{\pi}_{ij}(\tau)\, w_{ij}^2} \sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\} - \tau}$, where the $\widehat{\pi}_{ij}(\tau)$ are* *the empirical estimates of the q_i/Q_j computed at local time τ. Then, the regret satisfies for all $U \in \mathcal{U}$*

$$\mathbb{E}[R_T(U)] \overset{\widetilde{\mathcal{O}}}{=} \left(\sum_{j=1}^N \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} \sqrt{1 + \sigma_j^2(N_j - 1)} \right) \sqrt{T} + \frac{N}{q_{\min}} \,,$$

where the expectation is taken with respect to the agent activations, and $\widetilde{\mathcal{O}}$ neglects logarithmic terms in N and T.

Proof. We first introduce some notation. For any $j \in [N]$, $i \in \mathcal{N}_j$, and $t > 0$, let $t^{(j)} = \sum_{s \le t} \mathbb{I}\{i_s \in \mathcal{N}_j\}$ be the local time for the artificial clique centered at agent j, $\pi_{ij} = q_i/Q_j$, and $\widehat{\pi}_{ij}(t^{(j)}) = \sum_{s \le t} \mathbb{I}\{i_s = i\} / \sum_{s \le t} \mathbb{I}\{i_s \in \mathcal{N}_j\} = \sum_{s \le t} \mathbb{I}\{i_s = i\} / t^{(j)}$ be its natural estimator at local time $t^{(j)}$. The strategy is as follows. First, each agent j predicts with $Y_t^{(j)} = 0$ until its local time $t^{(j)}$ reaches some value τ_j to be determined later. It then builds the estimates $\widehat{\pi}_{ij}(\tau_j)$ of the π_{ij}. Finally, from local time τ_j agent j predicts $Y_t^{(j)}$ using its local instance of `MT-FTRL`, run with $N = N_j$, and $\beta_{t-1} = \sqrt{\sum_{i \in \mathcal{N}_j} \widehat{\pi}_{ij}(\tau_j) \, w_{ij}^2} \sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\} - \tau_j}$. We refer to this algorithm as `MT-FTRL-Ada`q.

Note that Lemma 2 still applies, such that we may only focus on the regret on the different instances of `MT-FTRL-Ada`q run by the different agents. We differentiate two cases, depending on whether the event

$$\mathcal{E} := \left\{ \frac{\pi_{ij}}{2} \le \widehat{\pi}_{ij}(\tau_j) \le \frac{3}{2}\pi_{ij} \,, \quad \forall j \in [N], \forall i \in \mathcal{N}_j \right\}$$

occurs or not. If $\mathcal{E}$ is not satisfied, the regret of the instance of `MT-FTRL-Ada`q run by agent j is trivially bounded by T_j, and that of the global approach by T. When $\mathcal{E}$ is satisfied, we can control the regret of the instance of `MT-FTRL-Ada`q run by agent j as follows. On the first τ_j round, the regret is bounded by τ_j. From time step τ_j onward, we have $\frac{\pi_{ij}}{2} \le \widehat{\pi}_{ij}(\tau_j) \le \frac{3}{2}\pi_{ij}$ and taking expectations on (17), with $\beta = \sqrt{\sum_{i \in \mathcal{N}_j} \widehat{\pi}_{ij}(\tau_j) \, w_{ij}^2}$, we obtain, conditionally on $\mathcal{E}$

$$\begin{aligned}
&\mathbb{E}\left[\sum_{\substack{t \colon t^{(j)} \ge \tau_j + 1 \\ t \colon i_t \in \mathcal{N}_j}} \left\langle w_{i_t j} \, g_t, \left[X_t\right]_{i_t :} - \left[X_t^{(\xi^*)}\right]_{i_t :} \right\rangle \,\middle|\, \mathcal{E} \right] \\
&\le \mathbb{E}\Bigg[\sqrt{\sum_{i \in \mathcal{N}_j} \widehat{\pi}_{ij}(\tau_j) \, w_{ij}^2} \sqrt{\ln N_j \sum_{t \colon t^{(j)} \ge \tau_j + 1} \mathbb{I}\{i_t \in \mathcal{N}_j\}} \\
&\qquad + \frac{\sqrt{\ln N_j}}{2\sqrt{\sum_{i \in \mathcal{N}_j} \widehat{\pi}_{ij}(\tau_j) \, w_{ij}^2}} \sum_{t \colon t^{(j)} = \tau_j + 1} \frac{w_{i_t j}^2 \, \mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{\substack{s \colon s^{(j)} \ge \tau_j + 1 \\ s \le t-1}} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \,\Bigg|\, \mathcal{E} \Bigg] \\
&\le \mathbb{E}\left[\sqrt{\frac{3}{2}} \sqrt{\sum_{i \in \mathcal{N}_j} \pi_{ij} \, w_{ij}^2} \sqrt{\ln N_j \sum_{t \ge 1} \mathbb{I}\{i_t \in \mathcal{N}_j\}} + \frac{\sqrt{2}\sqrt{\ln N_j}}{2\sqrt{\sum_{i \in \mathcal{N}_j} \pi_{ij} \, w_{ij}^2}} \sum_{t=1}^T \frac{w_{i_t j}^2 \, \mathbb{I}\{i_t \in \mathcal{N}_j\}}{\sqrt{1 + \sum_{s \le t-1} \mathbb{I}\{i_s \in \mathcal{N}_j\}}} \right] \\
&\le 2\sqrt{2} \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} \, \sqrt{T \, \ln N_j} \,.
\end{aligned}$$

The same method applies to the regret with choice ξ^*, see decomposition (9), and we finally obtain that the expected regret of the overall approach, is bounded by

$$\underbrace{\sum_{j=1}^N \tau_j}_{:= \tau} + \underbrace{9 \left(\sum_{j=1}^N \sqrt{\sum_{i \in \mathcal{N}_j} q_i \, w_{ij}^2} \sqrt{1 + \sigma_j^2(N_j - 1)} \right) \sqrt{T \, \ln N}}_{:= \mathfrak{A}} \,,$$

conditionally to $\mathcal{E}$. Hence, we have

$$\mathbb{E}[R_T(U)] \le \mathbb{E}\big[\mathbb{P}(\mathcal{E})(\tau + \mathfrak{A}) + \mathbb{P}(\mathcal{E}^{\mathsf{c}})T\big] \le \tau + \mathfrak{A} + \mathbb{E}\big[\mathbb{P}(\mathcal{E}^{\mathsf{c}})\big]T\,. \tag{18}$$

We now upper bound $\mathbb{P}(\mathcal{E}^{\mathsf{c}})$. Let $q_{\min} = \min_{i\in[N]} q_i$. For any $j \in [N]$ and $i \in \mathcal{N}_j$, by the multiplicative Chernoff bound (Mitzenmacher and Upfal, 2005, Corollary 4.6) we have

$$\mathbb{P}\left(\widehat{\pi}_{ij}(\tau_j) \le \frac{\pi_{ij}}{2} \text{ or } \widehat{\pi}_{ij}(\tau_j) \ge \frac{3}{2}\pi_{ij}\right) = \mathbb{P}\left(|\widehat{\pi}_{ij}(\tau_j) - \pi_{ij}| \ge \frac{\pi_{ij}}{2}\right) \le \exp\left(-\frac{\pi_{ij}}{12}\,\tau_j\right) \le \exp\left(-\frac{q_{\min}}{12}\,\tau_j\right),$$

and by the union bound

$$\mathbb{P}(\mathcal{E}^{\mathsf{c}}) \le N\sum_{j=1}^{N} \exp\left(-\frac{q_{\min}}{12}\,\tau_j\right). \tag{19}$$

Substituting (19) into (18), and setting $\tau_j = \left\lceil \frac{12}{q_{\min}} \ln\left(2N^2T\right)\right\rceil$ for all j, we get

$$\begin{aligned}
\mathbb{E}[R_T(U)] &\le \mathfrak{A} + \tau + 2NT\sum_{j=1}^{N}\exp\left(-\frac{q_{\min}}{12}\,\tau_j\right)\\
&\le 9\left(\sum_{j=1}^{N}\sqrt{\sum_{i\in\mathcal{N}_j} q_i\, w_{ij}^2}\sqrt{1+\sigma_j^2(N_j-1)}\right)\sqrt{T\,\ln N} + \frac{12N}{q_{\min}}\ln\left(2N^2T\right) + N + 1\\
&\le 9\left(\sum_{j=1}^{N}\sqrt{\sum_{i\in\mathcal{N}_j} q_i\, w_{ij}^2}\sqrt{1+\sigma_j^2(N_j-1)}\right)\sqrt{T\,\ln N} + \frac{14N}{q_{\min}}\ln\left(2N^2T\right).
\end{aligned}$$

□

A.5 Proof of Theorem 6

Theorem 6. *Let G be any communication graph. Then for any algorithm following Section 2's protocol:*

1. there exists a sequence of activations and gradients such that the algorithm suffers regret

$$\sup_{\substack{U\in\mathcal{U}\\ \sigma^2(U)\le\nu^2}} R_T(U) \ge \max\left\{\sqrt{1+\nu^2(N-1)},\, \sqrt{\alpha_2(G)}\right\}\frac{\sqrt{T}}{3}.$$

2. for any even number K, there exists a K-regular graph and a sequence of activations and gradients such that the algorithm suffers regret

$$\sup_{U\in\mathcal{U}} R_T(U) \ge \frac{1}{5}\sqrt{1+K\,\sigma_{\min}^2}\,\sqrt{\frac{NT}{K}}\,.$$

Proof. The first part of the lower bound is proved in (Cesa-Bianchi et al., 2022, Proposition 4). It establishes the existence of a comparator U and a sequence of activations and gradients such that for any algorithm we have

$$R_T(U) \ge \frac{\sqrt{2}}{4}\sqrt{1+\sigma^2(N-1)}\sqrt{T}\,. \tag{20}$$

Note that this lower bound is oblivious to the communication graph G, and applies in particular to algorithms that follow the protocol described in Section 2 (where information can only be exchanged along the edges of G). For the second part of the first lower bound, let $S_{\alpha_2}(G) = \{j_1, \dots, j_{\alpha_2}\}$ be a largest twice independent set of G. Due to the communication protocol, it is immediate to check that if only nodes in $S_{\alpha_2}(G)$ are activated, then no information from an active node can be transmitted to another active node. Hence, the regret incurred by the nodes in $S_{\alpha_2}(G)$ are independent. Now, we know from standard online learning lower bounds that there exists some $u_0 \in \mathbb{R}^d$ and a sequence $g_1, \dots, g_T$ such that the regret of a single agent is lower bounded by $\sqrt{T}/2$, see e.g., (Orabona, 2019, Theorem 5.1). Let $U = (u_0, \dots, u_0) \in \mathbb{R}^{N\times d}$, the activations be j_1 for the first $T/\alpha_2(G)$

Figure 4: Activated nodes in a 2-regular graph.

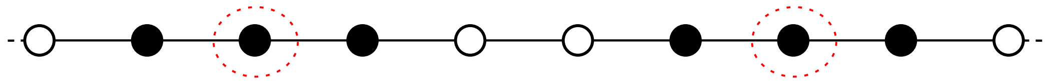


Figure 5: Activated nodes in a 4-regular graph.

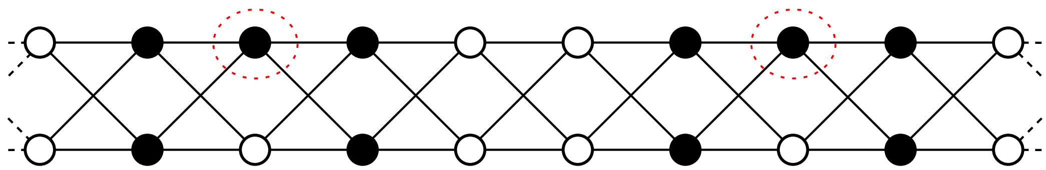
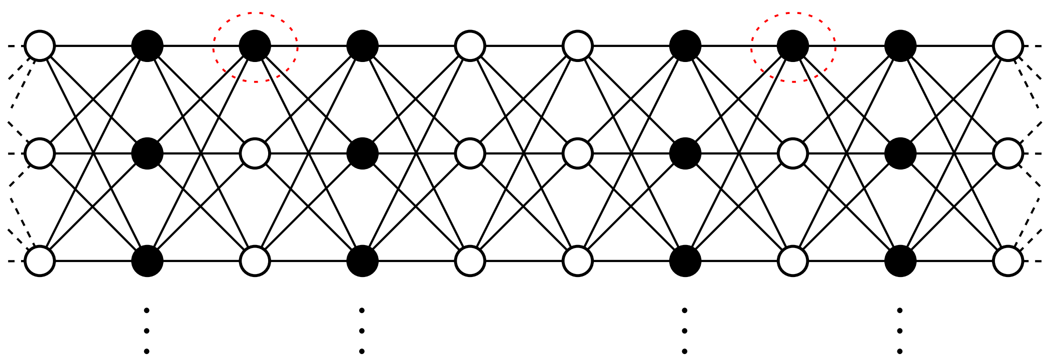


Figure 6: Activated nodes in a 6-regular graph and more.



steps, j_2 for the next $T/\alpha_2(G)$, and so on. Finally, consider the sequence of gradients $g_1 \dots g_{T/\alpha_2(G)}$ repeated $\alpha_2(G)$ times. In words, each agent in $S_{\alpha_2}(G)$ is sequentially given $g_1 \dots g_{T/\alpha_2(G)}$ and comparator u_0. Then the multitask regret satisfies

$$R_T(U) = \sum_{j \in S_{\alpha_2}(G)} R^{\text{agent } j}_{T/\alpha_2(G)} \geq \sum_{j \in S_{\alpha_2}(G)} \frac{\sqrt{T/\alpha_2(G)}}{2} = \frac{\sqrt{\alpha_2(G)\, T}}{2}\,, \tag{21}$$

where the first equality comes from the independence due to the communication constraints, and the inequality from the single-task lower bound. Gathering (20) and (21) yields the desired result.

Consider now the special case of K-regular graphs. Figures 4 to 6 show examples of communication graph on which our second lower bound can be attained. The idea is similar to that of the first lower bound. Only groups of nodes that are 2-nodes apart (the black nodes in Figures 4 to 6) are activated. This way, the regret incurred by the different groups are summed, as they are independent. On a group of nodes, one can identify a "central node" (circled with red dashes). We denote the set of indices of such nodes $\mathcal{J}_{\text{cn}}$. By the multitask lower bound of Cesa-Bianchi et al. (2022), we know that for the (artificial) clique with center node j there exists a comparator $U^{(j)} \in \mathbb{R}^{N_j \times d}$ and a sequence of activations and gradients such that we have a lower bound of order $\sqrt{1 + K\,\sigma_j^2}\sqrt{\sum_{i \in \mathcal{N}_j} T_i}$. Overall by feeding these gradients to each groups of nodes, each activated $T/|\mathcal{J}_{\text{cn}}|$ times, we obtain

$$R_T(U) = \sum_{j \in \mathcal{J}_{\text{cn}}} R^{\text{group centered at } j}_{T/|\mathcal{J}_{\text{cn}}|}(U^{(j)}) \geq \frac{\sqrt{2}}{4} \sum_{j \in \mathcal{J}_{\text{cn}}} \sqrt{1 + K\,\sigma_j^2}\,\sqrt{\frac{T}{|\mathcal{J}_{\text{cn}}|}} \geq \frac{1}{5}\sqrt{1 + K\,\sigma_{\min}^2}\,\sqrt{\frac{NT}{K}}\,,$$

where we have used that on the family of regular graphs depicted in Figures 4 to 6, we have $|\mathcal{J}_{\text{cn}}| = \frac{2N}{5K}$. $\square$

A.6 A Failing Approach

In MT-CO$_2$OL, each agent maintains an instance of MT-FTRL for the artificial clique composed of its neighbors. However, this is not the only way that one could use MT-FTRL as a building block to devise a generic algorithm operating on any communication graph. Another tempting approach consists in maintaining one instance of MT-FTRL for every pair of agents (i,j) such that $(i,j) \in E$. Let $Y_t^{(i,j)} \in \mathbb{R}^{2\times d}$ be the prediction maintained at time step t by such an algorithm.[4] Note that we have $Y_t^{(i,j)} = Y_t^{(j,i)}$ for any $(i,j) \in E$. Similarly to MT-CO$_2$OL, a natural way for the active agent i_t to make predictions is to submit the weighted average $x_t = \sum_{j\in\mathcal{N}_{i_t}} w_{i_t j}\big[Y_t^{(i_t,j)}\big]_{i_t:}$, where the w_{ij} are nonnegative and such that $\sum_{j\in\mathcal{N}_{i_t}} w_{ij} = 1$ for all i. Note that, contrary to MT-CO$_2$OL, predictions do not require a fetch step here, as both endpoints of each edge (i,j) are maintaining the same $Y_t^{(i,j)} = Y_t^{(j,i)}$. Using similar arguments as for MT-CO$_2$OL, we obtain

$$
\begin{aligned}
\sum_{t=1}^T \ell_t(x_t) - \ell_t(U_{i_t:}) &\le \sum_{t=1}^T \langle g_t, x_t - U_{i_t:}\rangle \\
&= \sum_{t=1}^T \sum_{i=1}^N \Big\langle g_t, \sum_{j\in\mathcal{N}_i} w_{ij}\big[Y_t^{(i,j)}\big]_{i:} - U_{i:}\Big\rangle \mathbb{I}\{i_t = i\} \\
&= \sum_{t=1}^T \sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \Big\langle g_t, w_{ij}\Big(\big[Y_t^{(i,j)}\big]_{i:} - U_{i:}\Big)\Big\rangle \mathbb{I}\{i_t = i\} \\
&= \sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \sum_{t=1}^T \Big\langle w_{ij}\, g_t, \Big(\big[Y_t^{(i,j)}\big]_{i:} - U_{i:}\Big)\Big\rangle \mathbb{I}\{i_t = i\} \qquad (22)\\
&= \sum_{j=1}^N \sum_{i\in\mathcal{N}_j} \sum_{t=1}^T \Big\langle w_{ij}\, g_t, \Big(\big[Y_t^{(i,j)}\big]_{i:} - U_{i:}\Big)\Big\rangle \mathbb{I}\{i_t = i\} \qquad (23)\\
&= \sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \sum_{t=1}^T \Big\langle w_{ji}\, g_t, \Big(\big[Y_t^{(i,j)}\big]_{j:} - U_{j:}\Big)\Big\rangle \mathbb{I}\{i_t = j\} \qquad (24)\\
&= \frac{1}{2}\sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \sum_{t=1}^T \Big[\Big\langle w_{ij}\, g_t, \Big(\big[Y_t^{(i,j)}\big]_{i:} - U_{i:}\Big)\Big\rangle \mathbb{I}\{i_t = i\} \\
&\qquad\qquad + \Big\langle w_{ji}\, g_t, \Big(\big[Y_t^{(i,j)}\big]_{j:} - U_{j:}\Big)\Big\rangle \mathbb{I}\{i_t = j\}\Big] \qquad (25)\\
&= \sum_{(i,j)\in E} R_T^{(i,j)}\big(U^{(i,j)}\big)\,,
\end{aligned}
$$

where (23) derives from Lemma 9, (24) is (23) with indices i and j swapped, (25) is the average of (22) and (24), and $R_T^{(i,j)}\big(U^{(i,j)}\big)$ denotes the regret suffered by the instance of MT-FTRL run by edge (i,j) on the linear losses $\big\langle (w_{ij}\mathbb{I}\{i_t = i\} + w_{ji}\mathbb{I}\{i_t = j\})\, g_t, \cdot\big\rangle$ over the rounds $t \le T$ such that $i_t = i$ or $i_t = j$. Now, by the MT-FTRL analysis, assuming stochastic activations we have

$$
R_T^{(i,j)}\big(U^{(i,j)}\big) \overset{\widetilde{\mathcal{O}}}{=} \sqrt{(q_i w_{ij}^2 + q_j w_{ji}^2)(1+\Delta_{ij}^2)}\sqrt{T}\,,
$$

where $\Delta_{ij}^2 = \|U_{i:} - U_{j:}\|_2^2$. Hence, if we assume uniform activations (i.e., $q_i = 1/N$ for all i) in the single-task case (i.e., $\Delta_{ij}^2 = 0$ for all $(i,j) \in E$) we obtain overall

$$
R_T(U) \overset{\widetilde{\mathcal{O}}}{=} \sum_{(i,j)\in E} \sqrt{w_{ij}^2 + w_{ji}^2}\sqrt{T/N}\,,
$$

[4] We consider here for simplicity that every agent is connected to itself, such that we also maintain $Y_t^{(i,i)}$ for all $i \in V$. The latter is exclusively based on the feedback received by agent i.

which can be further simplified into

$$\sum_{(i,j)\in E} \sqrt{w_{ij}^2 + w_{ji}^2}\sqrt{T/N} \leq \sum_{(i,j)\in E} \big(w_{ij} + w_{ji}\big)\sqrt{T/N} = \sqrt{NT}\,. \tag{26}$$

Contrary to $\texttt{MT-CO}_2\texttt{OL}$, this strategy thus fails to improve over the naive bound using independent updates (`iFTRL`). Note that the inequality in (26) is tight, while the equality holds for any w_{ij}. Hence, **there is no choice of weights that allows achieving better performance than independent runs of** `FTRL`. This can be intuitively explained as this strategy runs `MT-FTRL` on small subsets (pairs), which makes learning slower as they are activated less often. Instead, $\texttt{MT-CO}_2\texttt{OL}$ leverages `MT-FTRL` on the largest possible set of nodes that can communicate with the central agent, i.e., its neighbourhood.

B Technical Proofs (Results from Section 4)

In this section, we gather the technical proofs of the results exposed in Section 4. First, we provide some intuition about the construction of $\texttt{DPMT-CO}_2\texttt{OL}$. Then, we prove that $\texttt{DPMT-CO}_2\texttt{OL}$ is ε-DP (Theorem 7) before bounding its regret (Theorem 8).

B.1 Intuition on $\texttt{DPMT-CO}_2\texttt{OL}$ and Explicitation of `DPMT-FTRL`

In this section, we provide more intuition about the construction of $\texttt{DPMT-CO}_2\texttt{OL}$. In particular, we explain why $\texttt{DPMT-CO}_2\texttt{OL}$ invokes `DPMT-FTRL`, instead of any multitask algorithm working on a clique like $\texttt{MT-CO}_2\texttt{OL}$. `DPMT-FTRL` is a variant of `MT-FTRL` that works with a different kind of feedback. Instead of observing the gradient of the loss at the prediction, `DPMT-FTRL` observes sanitized versions of cumulative sums that are needed to compute the expert predictions $X_t^{(\xi)}$ and their weights $p_t^{(\xi)}$. To explain this difference, we dissect $\texttt{MT-CO}_2\texttt{OL}$ and highlight which changes are necessary to make it DP.

Extending the notation of Algorithms 1 and 4, we denote by $X_t^{(j,\xi)}$ and $p_t^{(j,\xi)}$ the experts and probabilities maintained by agent j. Let $\mathcal{U}_j := \big\{U \in \mathbb{R}^{N_j\times d} \colon U_{i:} \in \mathcal{X} \text{ for all } i \leq N_j\big\}$ and Δ_j the simplex in $\mathbb{R}^{N_j}$. In $\texttt{MT-CO}_2\texttt{OL}$, we have

$$\begin{aligned}
X_t^{(j,\xi)} &= \operatorname*{arg\,min}_{\substack{X\in\mathcal{U}_j\\ \sigma^2(X)\leq\xi}} \;\eta_{t-1}^{(\xi)} \sum_{s\leq t-1} \big\langle w_{i_s j}\, g_s, X_{i_s:}\big\rangle\, \mathbb{I}\left\{i_s \in \mathcal{N}_j\right\} + \frac{1}{2}\|X\|_{A_j}^2\\
&= \operatorname*{arg\,min}_{\substack{X\in\mathcal{U}_j\\ \sigma^2(X)\leq\xi}} \;\eta_{t-1}^{(\xi)} \sum_{s\leq t-1}\sum_{i\in\mathcal{N}_j} \big\langle w_{ij}\, g_s, X_{i:}\big\rangle\, \mathbb{I}\left\{i_s = i\right\} + \frac{1}{2}\|X\|_{A_j}^2\\
&= \operatorname*{arg\,min}_{\substack{X\in\mathcal{U}_j\\ \sigma^2(X)\leq\xi}} \;\eta_{t-1}^{(\xi)} \sum_{i\in\mathcal{N}_j} \Big\langle w_{ij} \underbrace{\sum_{s\leq t-1} g_s\, \mathbb{I}\left\{i_s = i\right\}}_{:=\gamma_t^{(i)}}, X_{i:}\Big\rangle + \frac{1}{2}\|X\|_{A_j}^2\,.
\end{aligned} \tag{27}$$

Hence, agent j must receive private versions of the $\gamma_t^{(i)}$ from each of its neighbor $i \in \mathcal{N}_j$, which are precisely the $\widetilde{\gamma}_t^{(i)}$ in Algorithm 3. Note that the $\gamma_t^{(i)}$ are sums of gradients, that can be sanitized efficiently using `TreeBasedAgg` (Chan et al., 2011). But the $\gamma_t^{(i)}$ are not the only gradient information agent j needs from its neighbors. Indeed, we have

$$\begin{aligned}
p_t^{(j,:)} &= \operatorname*{arg\,min}_{p\in\Delta_j} \;\frac{\sqrt{\ln N_j}}{\beta_{t-1}} \sum_{s\leq t-1} \langle \mathrm{loss}_s, p\rangle\, \mathbb{I}\left\{i_s \in \mathcal{N}_j\right\} + \sum_{\xi\in\Xi_j} p^{(\xi)} \ln p^{(\xi)}\\
&= \operatorname*{arg\,min}_{p\in\Delta_j} \;\frac{\sqrt{\ln N_j}}{\beta_{t-1}} \sum_{s\leq t-1}\sum_{\xi\in\Xi_j} p^{(\xi)} \Big\langle w_{i_s j}\, g_s, \big[X_s^{(j,\xi)}\big]_{i_s:}\Big\rangle \mathbb{I}\left\{i_s \in \mathcal{N}_j\right\} + \sum_{\xi\in\Xi_j} p^{(\xi)} \ln p^{(\xi)}\\
&= \operatorname*{arg\,min}_{p\in\Delta_j} \;\frac{\sqrt{\ln N_j}}{\beta_{t-1}} \sum_{\xi\in\Xi_j} p^{(\xi)} \sum_{i\in\mathcal{N}_j} w_{ij} \underbrace{\sum_{s\leq t-1} \Big\langle g_s, \big[X_s^{(j,\xi)}\big]_{i:}\Big\rangle \mathbb{I}\left\{i_s = i\right\}}_{:=s_{t,i}^{(j,\xi)}} + \sum_{\xi\in\Xi_j} p^{(\xi)} \ln p^{(\xi)}\,.
\end{aligned} \tag{28}$$

Algorithm 4 `DPMT-FTRL` (on linear losses)

Requires: Number of agents N, learning rates β_{t-1}

Init: $A = (1+N)I_N - \mathbf{1}_N \mathbf{1}_N^\top$, $\Xi = \{1/N, 2/N, \dots, 1\}$, $p_1^{(\xi)} = \frac{1}{N}\ \forall \xi \in \Xi$, $\widetilde{\Gamma}_0 = 0_{\mathbb{R}^{N\times d}}$, $\widetilde{s}_t^{(\xi)} = 0_{\mathbb{R}^N}\ \forall \xi \in \Xi$

for $t = 1, 2, \dots$ **do**

  for $\xi \in \Xi$ **do**

    `// Set learning rate assuming` $\sigma^2 = \xi$

    $\eta_{t-1}^{(\xi)} = \frac{N}{\beta_{t-1}}\sqrt{1+\xi(N-1)}$

    `// FTRL with Mahalanobis regularizer`

    $X_t^{(\xi)} = \underset{\substack{X\in\mathcal{U} \\ \sigma^2(X)\le\xi}}{\arg\min}\ \eta_{t-1}^{(\xi)}\left\langle \widetilde{\Gamma}_{t-1}, X\right\rangle + \dfrac{1}{2}\|X\|_A^2$

  `// Predict and receive feedback`

  Predict $Y_t = \sum_{\xi\in\Xi} p_t^{(\xi)} X_t^{(\xi)}$

  Incur loss $\langle g_t, [Y_t]_{i_t:}\rangle$ and receive

    $\widetilde{\gamma}_t^{(i_t)}$ the sanitized version of $\gamma_t^{(i_t)} = \sum_{s\le t} g_s \mathbb{I}\{i_s = i_t\}$

    $\widetilde{s}_{t,i_t}^{(\xi)}$ the sanitized version of $s_{t,i_t}^{(\xi)} = \sum_{s\le t-1}\left\langle g_s, \left[X_s^{(\xi)}\right]_{i_t:}\right\rangle \mathbb{I}\{i_s = i_t\}$, for all $\xi\in\Xi$

  `// Update` $\widetilde{\Gamma}_t$ `and` $\widetilde{s}_t^{(\xi)}$

  $[\widetilde{\Gamma}_t]_{i_t:} = \widetilde{\gamma}_t^{(i_t)}$ and $[\widetilde{\Gamma}_t]_{i:} = [\widetilde{\Gamma}_{t-1}]_{i:}\ \ \forall i \neq i_t$

  for $\xi \in \Xi$ **do**

    $\widetilde{s}_{t,i_t}^{(\xi)} = \widetilde{s}_{t,i_t}^{(\xi)}$ and $\widetilde{s}_{t,i}^{(\xi)} = \widetilde{s}_{t-1,i}^{(\xi)}\ \ \forall i \neq i_t$

  `// Update` p_t `based on the experts losses`

  for $\xi \in \Xi$ **do**

    $p_{t+1}^{(\xi)} = \dfrac{\exp\left(-\frac{\sqrt{\ln N}}{\beta_t}\sum_{i=1}^N \widetilde{s}_{t,i}^{(\xi)}\right)}{\sum_k \exp\left(-\frac{\sqrt{\ln N}}{\beta_t}\sum_{i=1}^N \widetilde{s}_{t,i}^{(k)}\right)}$

Hence, agent j also needs to receive private versions of the $s_{t,i}^{(j,\xi)}$, which are the $\widetilde{s}_{t,i}^{(j,\xi)}$ in Algorithm 3. Note that, again, the $s_{t,i}^{(j,\xi)}$ are sums, so that they can be sanitized easily using `TreeBasedAgg`.

The tree aggregations needed to sanitize the $\gamma_t^{(i)}$ and the $s_{t,i}^{(j,\xi)}$ are taken care of by instructions contained in Algorithm 3, while the modifications of `MT-FTRL` needed to take private cumulative sums as inputs are outlined in Algorithm 4. In the latter pseudo-code, we drop the superscripts $^{(j)}$ since we consider `DPMT-FTRL` on a clique only.

Information flow. The above discussion allows understanding which information needs to be exchanged between neighbors. In particular, the messages exchanged between agents, even if sanitized, carry information about the gradient sequence. Consequently, analyzing the content of these messages is crucial for the analysis in terms of DP of `DPMT-CO`$_2$`OL`. Overall, the message $m_t^{(i\to j)}$ sent by agent i to agent j at iteration t of `DPMT-CO`$_2$`OL` is

$$m_t^{(i\to j)} = \begin{cases} p_t^{(i,:)},\ \left\{\left[X_t^{(i,\xi)}\right]_{j:} : \xi\in\Xi_i\right\} & \text{if } i\in\mathcal{N}_{i_t}\setminus\{i_t\} \text{ and } j = i_t \text{ (fetch step)} \\ w_{ij}\,\widetilde{\gamma}_t^{(i)},\ \left\{\widetilde{s}_{t,i}^{(j,\xi)} : \xi\in\Xi_j\right\} & \text{if } i = i_t \text{ and } j\in\mathcal{N}_{i_t} \text{ (send step)} \\ \emptyset & \text{otherwise} \end{cases} \tag{29}$$

Note that sending $p_t^{(i,:)}$ and $\left\{\left[X_t^{(i,\xi)}\right]_{j:} : \xi\in\Xi_i\right\}$ is actually equivalent (information and privacy-wise) to sending $\left[Y_t^{(i)}\right]_{i_t:}$ as written in Algorithm 3, since agent i_t can compute $\left[Y_t^{(i)}\right]_{i_t:} = \sum_{\xi\in\Xi_i} p_t^{(i,\xi)}\left[X_t^{(i,\xi)}\right]_{i_t:}$. In addition, having access to the individual experts $\left[X_t^{(i,\xi)}\right]_{j:}$ is necessary for i_t to compute the $\widetilde{s}_{t,i_t}^{(i,\xi)}$. In what follows we denote $m_t^{(i)} = \{m_t^{(i\to j)} \colon j \in \mathcal{N}_i\}$ the batch of messages sent by agent i to other agents at time step t.

B.2 Proof of Theorem 7

Theorem 7. *Let G be any graph, and for any $\varepsilon > 0$, let $\varepsilon' = \varepsilon/(6N_{\max}^2)$. Assume that* `DPMT-CO2OL` *is run with $\mathcal{D}_d$ set to a d-dimensional Laplacian distribution with parameter $\frac{\sqrt{d}\ln T}{\varepsilon'}$, and $\mathcal{D}_1$ set to a 1-dimensional Laplacian distribution with parameter $\frac{\ln T}{\varepsilon'}$. Then* `DPMT-CO2OL` *is ε-DP.*

Proof. Recall from Definition 2 that we should check that for any $i \in [N]$ and set of message sequences $\mathcal{M}$ we have

$$\frac{\mathbb{P}\big(m_1^{(i)},\dots,m_T^{(i)} \in \mathcal{M} \mid i_1,g_1,\dots,i_T,g_T\big)}{\mathbb{P}\big(m_1^{(i)},\dots,m_T^{(i)} \in \mathcal{M} \mid i_1,g_1',\dots,i_T,g_T'\big)} \le e^{\varepsilon}\,.$$

We will show that the above equation holds by focusing on each possible component of message $m_t^{(i)}$, see (29). Let τ be the round where sequences of gradients $(g_t)_{t\in[T]}$ and $(g_t')_{t\in[T]}$ differ, we start by identifying which messages are impacted by the change of g_τ into g_τ':

- By definition, $\widetilde{\gamma}_t^{(i)}$ is impacted iff $i = i_\tau$.
- Recalling (27), $X_t^{(i,\xi)}$ is impacted iff $i \in \mathcal{N}_{i_\tau}$.
- By definition, $\widetilde{s}_{t,i}^{(j,\xi)}$ is impacted iff $\sum_t g_t \mathbb{I}\{i_t = i\}$ or the $X_t^{(j,\xi)}$ are impacted, i.e., iff $i = i_t$ or $j \in \mathcal{N}_{i_\tau}$.
- Recalling (28), $p_t^{(i,:)}$ is impacted iff any $\widetilde{s}_{t,j}^{(i,\xi)}$ for $j \in \mathcal{N}_i$ is impacted, i.e., iff $i \in \mathcal{N}_{i_\tau}$.

We now quantify the ratios of probabilities that the above elements of messages belong to $\mathcal{M}$, given the different sequences of gradients. Note that the latter elements may not belong to the same space (e.g., $p_t^{(i,:)} \in \Delta_i$ while $\widetilde{\gamma}_t^{(i)} \in \mathbb{R}^d$), such that in the following we use $\mathcal{M}$ as a generic set that adapts to the type of element considered. We highlight that for any i,j such that the message element is not impacted, the probability ratio is equal to 1, such that in particular it is bounded by the terms exhibited at Equations (30), (31), (33) and (34).

Recall that variables $\widetilde{\gamma}_t^{(i)}$ are generated by a tree aggregation with at most T summands, that have a maximal L_1 norm of $\sqrt{d}\max_{i\in\mathcal{N}_j} w_{ij} \le \sqrt{d}$, and a noise following a Laplacian distribution with parameter $\sqrt{d}\frac{\ln T}{\varepsilon'}$. Hence by Chan et al. (2011, Theorem 3.5) and the decomposition introduced in Smith et al. (2017, Equation (11)), allowing to deal with the dependence on a given gradient at some round of subsequent gradients, for any $i \in [N]$ we have

$$\frac{\mathbb{P}\big(\widetilde{\gamma}_1^{(i)},\dots,\widetilde{\gamma}_T^{(i)} \in \mathcal{M} \mid i_1,g_1,\dots,i_T,g_T\big)}{\mathbb{P}\big(\widetilde{\gamma}_1^{(i)},\dots,\widetilde{\gamma}_T^{(i)} \in \mathcal{M} \mid i_1,g_1',\dots,i_T,g_T'\big)} \le \exp(\varepsilon')\,. \tag{30}$$

We remind that the noise completion introduced by Agarwal and Singh (2017) does not harm the privacy, as it can be considered as post-processing.

Next, let $X_t^{(i,:)} = \{X_t^{(i,\xi)} \colon \xi \in \Xi_i\}$. By the construction of the $X_t^{(i,\xi)}$, which are deterministic functions of the $\gamma_t^{(j)}$ for $j \in \mathcal{N}_i$, see (27), we have for any $i \in [N]$

$$\frac{\mathbb{P}\big(X_1^{(i,:)},\dots,X_T^{(i,:)} \in \mathcal{M} \mid i_1,g_1,\dots,i_T,g_T\big)}{\mathbb{P}\big(X_1^{(i,:)},\dots,X_T^{(i,:)} \in \mathcal{M} \mid i_1,g_1',\dots,i_T,g_T'\big)} \le \exp(\varepsilon')\,. \tag{31}$$

We now focus on the $\widetilde{s}_{t,i}^{(j,\xi)}$, and denote $\widetilde{s}_{t,i}^{(j,:)} = \{\widetilde{s}_{t,i}^{(j,\xi)} \colon \xi \in \Xi_j\}$. For any i,j we have

$$\frac{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,:)},\dots,\widetilde{s}_{T,i}^{(j,:)} \in \mathcal{M} \mid i_1,g_1,\dots,i_T,g_T\Big)}{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,:)},\dots,\widetilde{s}_{T,i}^{(j,:)} \in \mathcal{M} \mid i_1,g_1',\dots,i_T,g_T'\Big)}$$

$$
\begin{aligned}
&= \int_{\mathcal{X}} \frac{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,:)}, \dots, \widetilde{s}_{T,i}^{(j,:)} \in \mathcal{M} \mid i_1, g_1, \dots, i_T, g_T, \big(X_s^{(j,:)}\big)_{s\in[T]} \in \partial\mu\Big)\, \mathbb{P}\Big(\big(X_s^{(j,:)}\big)_{s\in[T]} \in \partial\mu \mid i_1, g_1, \dots, i_T, g_T\Big)}{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,:)}, \dots, \widetilde{s}_{T,i}^{(j,:)} \in \mathcal{M} \mid i_1, g'_1, \dots, i_T, g'_T, \big(X_s^{(j,:)}\big)_{s\in[T]} \in \partial\mu\Big)\, \mathbb{P}\Big(\big(X_s^{(j,:)}\big)_{s\in[T]} \in \partial\mu \mid i_1, g'_1, \dots, i_T, g'_T\Big)} \partial\mu \\
&\le \prod\nolimits_{\xi\in\Xi_j} \int_{\mathcal{X}'} \frac{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,\xi)}, \dots, \widetilde{s}_{T,i}^{(j,\xi)} \in \mathcal{M} \mid i_1, g_1, \dots, i_T, g_T, \big(X_s^{(j,\xi)}\big)_{s\in[T]} \in \partial\mu'\Big)}{\mathbb{P}\Big(\widetilde{s}_{1,i}^{(j,\xi)}, \dots, \widetilde{s}_{T,i}^{(j,\xi)} \in \mathcal{M} \mid i_1, g'_1, \dots, i_T, g'_T, \big(X_s^{(j,\xi)}\big)_{s\in[T]} \in \partial\mu'\Big)} \times \exp(\varepsilon')\,\partial\mu' \qquad (32) \\
&\le \exp\big((N_j+1)\varepsilon'\big) \qquad (33) \\
&\le \exp\big(2N_j\varepsilon'\big)\,,
\end{aligned}
$$

where (32) derives from (31), and (33) from Chan et al. (2011, Theorem 3.5) and the decomposition introduced in Smith et al. (2017, Equation (11)), since each $\widetilde{s}_{t,i}^{(j,\xi)}$ is generated by a tree aggregation with at most T inner products between the predictions of the experts and the gradients (which in particular have a L_1 norm bounded by $\max_{j\in\mathcal{N}_i} w_{ij} \le 1$), and a noise following a Laplacian distribution with parameter $\ln T/\varepsilon'$. $\mathcal{X}$ denotes the space of sequences of real matrices of dimension $N_j \times N_j$ $\mathcal{X}'$ denotes the space of sequences of vectors of dimension N_j.

Finally, noticing that $p_t^{(i,:)}$ is just a post-processing of the $\widetilde{s}_{t,j}^{(i,\xi)}$ for $j \in \mathcal{N}_i$ and $\xi \in \Xi_i$, see (28), we have for any $i \in [N]$

$$
\frac{\mathbb{P}\big(p_1^{(i,:)}, \dots, p_T^{(i,:)} \in \mathcal{M} \mid i_1, g_1, \dots, i_T, g_T\big)}{\mathbb{P}\big(p_1^{(i,:)}, \dots, p_T^{(i,:)} \in \mathcal{M} \mid i_1, g'_1, \dots, i_T, g'_T\big)} \le \exp\big(N_i(N_i+1)\varepsilon'\big) \le \exp\big(2N_i^2\varepsilon'\big)\,. \qquad (34)
$$

Overall, $\big(m_t^{(i)}\big)_{t\in[T]}$ is a post-processing of: $\big(\widetilde{\gamma}_t^{(i)}\big)_{t\in[T]}$, $\big(X_t^{(i,:)}\big)_{t\in[T]}$, $\big(\widetilde{s}_{t,i}^{(j,:)}\big)_{t\in[T], j\in\mathcal{N}_i}$, $\big(p_t^{(i,:)}\big)_{t\in[T]}$, such that for all $i \in [N]$ we have

$$
\frac{\mathbb{P}\big(m_1^{(i)}, \dots, m_T^{(i)} \in \mathcal{M} \mid i_1, g_1, \dots, i_T, g_T\big)}{\mathbb{P}\big(m_1^{(i)}, \dots, m_T^{(i)} \in \mathcal{M} \mid i_1, g'_1, \dots, i_T, g'_T\big)} \le \exp\big(\varepsilon' + \varepsilon' + 2N_i^2\varepsilon' + 2N_i^2\varepsilon'\big) \le \exp\big(6N_{\max}^2\varepsilon'\big) = \exp(\varepsilon)\,,
$$

where we have used that $\varepsilon' = \varepsilon/(6N_{\max}^2)$.

□

B.3 Proof of Theorem 8

We first recall two results, which are instrumental to our analysis.

Lemma 11 (An alternative version of Theorem 3.4 of Agarwal and Singh (2017))**.** *For any noise distribution $\mathcal{D}$, regularizers $\psi_t(x) = \frac{\psi(x)}{\eta_{t-1}}$ with non decreasing $(\eta_t)_{t\ge1}$ and ψ that is μ-strongly convex with respect to $\|\cdot\|$, decision set $\mathcal{X}$, and loss vectors $g_1, \dots g_T$, the regret of* `FTRL` *on sums of gradients sanitized through* `TreeBasedAgg` *set with distribution $\mathcal{D}$ is bounded by*

$$
\mathbb{E}[\mathrm{Regret}_T(u)] \le \frac{\psi(u) - \min_{x\in\mathcal{X}} \psi(x)}{\eta_T} + \frac{1}{2\mu}\sum_{t=1}^{T} \eta_{t-1}\|g_t\|_*^2 + D_{\mathcal{D}'}\,,
$$

where $D_{\mathcal{D}'} = \mathbb{E}_{Z\sim\mathcal{D}'}\big[\max_{x\in\mathcal{X}}\langle Z, x\rangle - \min_{x\in\mathcal{X}}\langle Z, x\rangle\big]$, and $\mathcal{D}'$ is the probability distribution of the sum of $\ln T$ variables with distribution $\mathcal{D}$.

Proof. This is an alternative version of Theorem 3.4 of Agarwal and Singh (2017), but with non-constant learning rates. The proof relies on the following observation, already proved by Agarwal and Singh (2017). Let $(x_t)_{t\in[T]}$ denote the predictions of the instance of `FTRL` considered in Lemma 11. Consider the alternative experiment in which

$$
\begin{cases}
Z \text{ is such that } Z \sim \sum_{k\in[\ln T]} Z_k \text{ where } Z_k \sim \mathcal{D}, \\
\widehat{x}_1 = x_1, \\
\widehat{x}_{t+1} = \arg\min_{X\in\mathcal{X}} \ \big\langle \sum_{s=1}^{t} g_s + Z, x\big\rangle + \psi_t(x),
\end{cases}
$$

Then we have

$$\mathbb{E}[\mathrm{Regret}_T(u)] = \mathbb{E}\big[\widehat{\mathrm{Regret}}_T(u)\big]\,, \tag{35}$$

where $\widehat{\mathrm{Regret}}_T$ is regret of the strategy outputting $\widehat{x}_1, \dots \widehat{x}_T$. The complete proof of Equation (35) can be found in Agarwal and Singh (2017), but we provide its two main arguments. First, since the losses are linear, subsequent gradients do not depend on x_t. Second, $\sum_{s=1}^t g_s + Z$ is distributed as the sanitized sum of gradients used in `FTRL`.

Equation (35) proves that the expected regret is equal to the expectation of $\widehat{\mathrm{Regret}}_T$, the regret of `FTRL` with an alternative regularizer $\psi_Z(\cdot) = \psi(\cdot) + \langle Z, \cdot\rangle$. The addition of $\langle Z, \cdot\rangle$ does not change the strongly convex nature of the regularizer, since it is linear. The usual analysis of `FTRL` with ψ_Z (Orabona, 2019, Chapter 7) yields

$$\begin{aligned}
\mathbb{E}\big[\widehat{\mathrm{Regret}}_T(u)\big] &\le \frac{\psi_Z(u) - \min_{x\in\mathcal{X}} \psi_Z(x)}{\eta_T} + \frac{1}{2\mu}\sum_{t=1}^T \eta_{t-1}\|g_t\|_*^2 \\
&\le \frac{\psi(u) - \min_{x\in\mathcal{X}} \psi(x)}{\eta_T} + \frac{1}{2\mu}\sum_{t=1}^T \eta_{t-1}\|g_t\|_*^2 + D_{\mathcal{D}'}\,.
\end{aligned}$$

Using this jointly with Equation (35) concludes the proof. □

Lemma 12. *Let $(p_t)_{t\in[T]}$ be the sequence of probabilities maintained by the instance of `MT-FTRL` run by agent j among `DPMT-CO`$_2$`OL`, see Algorithm 3. Consider the alternative experiment in which*

$$\begin{cases} Z \text{ is such that } Z \sim \sum_{i\in\mathcal{N}_j}\sum_{k\in[\ln T]} Z_k, \text{ where } Z_k \sim \mathcal{D}_j, \text{ a Laplacian of dimension } N_j \text{ and parameter } \frac{\ln T}{\varepsilon'} \\ \widehat{p}_1 = p_1, \\ \widehat{p}_{t+1} = \arg\min_{p\in\Delta_j} \left\langle \sum_{s=1}^t \mathrm{loss}_s + Z, p\right\rangle + \sum_{\xi\in\Xi_j} p^{(\xi)} \ln p^{(\xi)} \end{cases}$$

Then for all $u \in \Delta_j$ we have

$$\mathbb{E}[\mathrm{Regret}_T(u)] = \mathbb{E}\big[\widehat{\mathrm{Regret}}_T(u)\big]\,,$$

where Regret_T and $\widehat{\mathrm{Regret}}_T$ are the regrets of the strategy outputting $p_1, \dots p_T$ and $\widehat{p}_1, \dots \widehat{p}_T$ respectively.

Proof. First observe that the probabilities p_t are the result of running `FTRL` where, in place of gradient inputs, a vector whose ξ-th element is $\sum_{i\in\mathcal{N}_j} \widetilde{s}_{t,i}^{(j,\xi)}$ is used. Note that the $\widetilde{s}_{t,i}^{(j,\xi)}$ come from instances of `TreeBasedAgg` set with distribution $\mathcal{D}_1$, so that the element ξ of the vector $\sum_{s=1}^t \mathrm{loss}_s + Z$ is actually distributed exactly as $\widetilde{s}_{t,i}^{(j,\xi)}$. This eventually shows that $\widehat{p}_t$ is distributed as p_t. □

We are now ready to prove Theorem 8.

Theorem 8. *Let $G = (V, E)$ be any communication graph, Consider `DPMT-CO`$_2$`OL` where the base algorithm run by each agent $j \in V$ is an instance of `DPMT-FTRL` with parameters $N = N_j$ and $\beta_{t-1} = \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1 + \sum_{s\le t-1}\mathbb{I}\{i_s \in \mathcal{N}_j\}}$. Then the regret of `DPMT-CO`$_2$`OL` with $\mathcal{D}_d$ and $\mathcal{D}_1$ set as in Theorem 7 satisfies*

$$\begin{aligned}
\mathbb{E}[R_T(U)] &\overset{\widetilde{\mathcal{O}}}{=} \sum_{j=1}^N \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1 + \sigma_j^2(N_j - 1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i} \\
&\quad + \frac{dN_{\max}^4}{\varepsilon} N \ln^2 T\,,
\end{aligned} \tag{7}$$

where randomness is due to sanitization.

Proof. Let $R_T^{\text{clique-}j}$ be the regret suffered by `MT-FTRL` run by agent j on the linear losses $\langle w_{i_t j}\, g_t, \cdot\rangle$, with feedback equal to $w_{i_t j}\, \widetilde{\gamma}_t^{(i_t)}$ for the sum of gradients incurred by i_t, and $\widetilde{s}_{t,i_t}^{(j,\xi)}$ for the sums of dot products between experts

(j, ξ) and the sum of gradients incurred by i_t. With the same steps used to prove Lemma 2, we can prove that the regret of `DPMT-CO`$_2$`OL` satisfies

$$\forall U \in \mathbb{R}^{N\times d}, \quad R_T(U) \le \sum_{j=1}^{N} R_T^{\text{clique-}j}\big(U^{(j)}\big)\,.$$

By the same arguments as in the proof of Theorem 1 we have

$$R_T^{\text{clique-}j}\big(U^{(j)}\big) \le \underbrace{\sum_{t\colon i_t\in\mathcal{N}_j} \Big\langle w_{i_tj}\, g_t, \big[Y_t^{(j)}\big]_{i_t:} - \big[X_t^{(j,\xi^*)}\big]_{i_t:}\Big\rangle}_{\text{Regret of } \texttt{Hedge}} + \underbrace{\sum_{t\colon i_t\in\mathcal{N}_j} \Big\langle w_{i_tj}\, g_t, \big[X_t^{(j,\xi^*)}\big]_{i_t:} - U_{i_t:}\Big\rangle}_{\text{Regret with choice } \xi^*},$$

where $\xi^* = \arg\min_{\xi\in\Xi_j} \sum_{t\colon i_t\in\mathcal{N}_j} \Big\langle w_{i_tj}\, g_t, \big[X_t^{(j,\xi)}\big]_{i_t:}\Big\rangle$. We start by upper bounding the regret due to `Hedge`. Let $\text{loss}_t \in \mathbb{R}^{N_j}$ storing the $w_{i_tj}\big\langle g_t, \big[X_t^{(j,\xi)}\big]_{i_t:}\big\rangle$ for $\xi \in \Xi_j$, and $e^* \in \mathbb{R}^{N_j}$ the one-hot vector with an entry of 1 at expert ξ^*. By the analysis of `Hedge` with regularizers $\psi_t(\boldsymbol{p}) = \frac{\beta\sqrt{1+\sum_{s\le t-1}\mathbb{I}\{i_s\in\mathcal{N}_j\}}}{\sqrt{\ln N_j}} \sum_{k=1}^{N_j} p_k \ln(p_k)$, combined with Lemma 12,

$$\begin{aligned}
\mathbb{E}\left[\sum_{t\colon i_t\in\mathcal{N}_j} \Big\langle w_{i_tj}\, g_t, \big[Y_t^{(j)}\big]_{i_t:} - \big[X_t^{(j,\xi^*)}\big]_{i_t:}\Big\rangle\right] &= \mathbb{E}\left[\sum_{t\colon i_t\in\mathcal{N}_j} \big\langle \text{loss}_t, \boldsymbol{p}_t - e^*\big\rangle\right] + D_{\mathcal{D}'_{1,j}} \\
&= 2\max_{i\in\mathcal{N}_j} w_{ij} \sqrt{\ln N_j \sum_{i\in\mathcal{N}_j} T_i} + D_{\mathcal{D}'_{1,j}}\,, \qquad (36)
\end{aligned}$$

where $D_{\mathcal{D}'_{1,j}} = \mathbb{E}_{Z\sim\mathcal{D}'_{1,j}}\big[\max_{x\in\Delta_j}\langle Z, x\rangle - \min_{x\in\Delta_j}\langle Z, x\rangle\big]$ and $\mathcal{D}'_{1,j}$ is the probability distribution of the sum of $N_j \ln T$ variables of distribution $\mathcal{D}_j$, which is a Laplacian of dimension N_j and parameter $\frac{\ln T}{\varepsilon'}$. Hence we have

$$\begin{aligned}
D_{\mathcal{D}'_{1,j}} &\le \mathbb{E}_{Z\sim\mathcal{D}'_{1,j}}\left[\max_{p\in\Delta_j}\langle Z, p\rangle - \min_{p\in\Delta_j}\langle Z, p\rangle\right] \\
&\le 2\max_{p\in\Delta_j}\|p\|_1 \mathbb{E}_{Z\sim\mathcal{D}'_{1,j}}\|Z\|_\infty \\
&\le 2\mathbb{E}_{Z\sim\mathcal{D}'_{1,j}}\|Z\|_\infty \\
&\le 2N_j \ln T\ \mathbb{E}_{Z\sim\mathcal{D}_j}\|Z\|_\infty \\
&\le 2N_j^{3/2} \ln T\ \mathbb{E}_{Z\sim\mathcal{D}_j}\|Z\|_2 \qquad (37) \\
&\le 2N_j^{3/2} \ln T \sqrt{\mathbb{E}_{Z_i\sim\mathcal{D}_1}\left[\sum_{i\in[N_j]} Z_i^2\right]} \\
&\le 2N_j^{2} \ln T\ \sqrt{\mathbb{E}_{Z\sim\mathcal{D}_1}\left[Z^2\right]} \\
&\le 2N_j^2 \frac{\ln^2 T}{\varepsilon'}\,, \qquad (38)
\end{aligned}$$

where (37) stems form the fact that Z is a vector of dimension N_j and (38) stems from the formula of the variance of a Laplacian random variable.

Now, we turn to the regret with choice ξ^*. Assume first that $\sigma_j^2 \le 1$. Then, the regret with choice ξ^* is in particular better than the regret with choice $\bar{\xi} \in \Xi_j$ such that $\bar{\xi} - \frac{1}{N_j} \le \sigma_j^2 \le \bar{\xi}$. Recall that the sequence $X_t^{(\bar{\xi})}$ is generated by `FTRL` with the sequence of regularizers $\frac{1}{2}\|\cdot\|_{A_j}^2/\eta_{t-1}^{(\bar{\xi})}$. By the analysis of `FTRL` and Lemma 11 applied to non-adaptive `MT-FTRL`, we retrieve

$$\mathbb{E}\left[\sum_{t\colon i_t\in\mathcal{N}_j} \Big\langle w_{i_tj}\, g_t, \big[X_t^{(j,\bar{\xi})}\big]_{i_t:} - U_{i_t:}\Big\rangle\right] \le 4\max_{i\in\mathcal{N}_j} w_{ij} \sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i} + D_{\mathcal{D}'_{d,j}}\,, \qquad (39)$$

where $D_{\mathcal{D}'_{d,j}} = \mathbb{E}_{Z\sim\mathcal{D}'_d}\big[\max_{x\in\mathcal{X}}\langle Z,x\rangle - \min_{x\in\mathcal{X}}\langle Z,x\rangle\big]$ and $\mathcal{D}'_{d,j}$ is the probability distribution of $\mathbb{R}^{N_j\times d}$ matrices whose rows are the sum of $\ln T$ variables of distribution $\mathcal{D}_d$, which is a Laplacian of dimension d and parameter $\frac{\sqrt{d}\ln T}{\varepsilon'}$. Now let us bound the additional term due to the privacy. We have

$$\begin{aligned} D_{\mathcal{D}'_{d,j}} &= \mathbb{E}\left[\max_{X\in\mathcal{X}}\langle Z,X\rangle - \min_{X\in\mathcal{X}}\langle Z,X\rangle\right] \le 2\mathbb{E}\left[\|Z\|_{A_j^{-1}}\ \max_{X\in\mathcal{X}}\|X\|_{A_j}\right] \\ &\le 2\sqrt{N_j}\sqrt{1+\bar{\xi}(N_j-1)}\ \mathbb{E}\|Z\|_{A_j^{-1}} \le 2\sqrt{2N_j}\sqrt{1+\sigma_j^2(N_j-1)}\ \mathbb{E}\|Z\|_{A_j^{-1}}\,. \end{aligned} \tag{40}$$

Now, we have

$$\begin{aligned} \mathbb{E}\|Z\|_{A_j^{-1}} &\le \sqrt{\mathbb{E}\|Z\|^2_{A_j^{-1}}} \\ &= \sqrt{\sum_{i,k\in[N_j]}[A_j^{-1}]_{ik}\ \mathbb{E}\left[ZZ^\top\right]_{ik}} \\ &= \sqrt{\sum_{i\in[N_j]}[A_j^{-1}]_{ii}\ \mathbb{E}\left[Z_{i:}^\top Z_{i:}\right]} \\ &\le \sqrt{\sum_{i\in[N_j]}[A_j^{-1}]_{ii}}\ \sqrt{\sum_{k=1}^d\mathbb{E}\left[Z_{1,k}^2\right]} \\ &\le \sqrt{\frac{2N_j}{1+N_j}}\sqrt{d}\sqrt{\mathrm{Var}\left[Z_{1,1}\right]} \\ &\le \sqrt{2d}\sqrt{\ln T\left(\frac{\sqrt{d}\ln T}{\varepsilon'}\right)^2} \\ &\le 2d\,\frac{\ln^2 T}{\varepsilon'}\,, \end{aligned}$$

where the third inequality stems from $[A_j{}^{-1}]_{ii} = \frac{2}{N_j+1}$ (see for example computations in Appendix A.2 of Cesa-Bianchi et al. (2022)). This in turn implies

$$D_{\mathcal{D}'_{d,j}} \le 4d\sqrt{2N_j}\sqrt{1+\sigma_j^2(N_j-1)}\,\frac{\ln^2 T}{\varepsilon'}\,. \tag{41}$$

Assume now that $\sigma_j^2 \ge 1$. Then, the regret with choice ξ^* is in particular better than the regret with choice 1. The latter corresponds to independent learning (Cesa-Bianchi et al., 2022) and an analysis similar to the one above shows that its regret is bounded by

$$\begin{aligned} \mathbb{E}\left[\sum_{t:\ i_t\in\mathcal{N}_j}\left\langle w_{i_tj}\,g_t,\left[X_t^{(j,\bar{\xi})}\right]_{i_t:}-U_{i_t:}\right\rangle\right] &\le \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{N_j\sum_{i\in\mathcal{N}_j}T_i} + D_{\mathcal{D}'_{d,j}} \\ &\le \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j}T_i} + D_{\mathcal{D}'_{d,j}}\,. \end{aligned} \tag{42}$$

In this case, we have

$$D_{\mathcal{D}'_{d,j}} \le 4dN_j\,\frac{\ln^2 T}{\varepsilon'} \le 4d\sqrt{2N_j}\sqrt{1+\sigma_j^2(N_j-1)}\,\frac{\ln^2 T}{\varepsilon'}\,. \tag{43}$$

Substituting (38) into (36) and (41) into (39) (or (43) into (42), depending on the value of σ_j^2), we obtain

$$\begin{aligned} &\mathbb{E}[R_T(U)] \\ &\quad\le \sum_{j=1}^N\left[6\max_{i\in\mathcal{N}_j} w_{ij}\left(\sqrt{1+\sigma_j^2(N_j-1)}+\ln N_j\right)\sqrt{\sum_{i\in\mathcal{N}_j}T_i} + D_{\mathcal{D}'_{1,j}} + D_{\mathcal{D}'_{d,j}}\right] \end{aligned}$$

Algorithm 5 `MT-FTRL` (with KT adaptive learning rate)

Requires: Learning rates β_t

Init: $A = (1+N)I_N - \mathbf{1}_N\mathbf{1}_N^\top$, $z_1 = 0$

for $t = 1, 2, \dots$ **do**

`// Update the direction using FTRL with Mahalanobis regularizer`

$$\widetilde{Y}_t = \underset{X\colon \|X\|_A \le 1}{\arg\min} \; \frac{\sqrt{N}}{\beta_{t-1}} \left\langle \sum_{s=1}^{t-1} G_s, X \right\rangle + \frac{1}{2}\|X\|_A^2$$

`// Update the magnitude`

$$z_t = -\frac{1}{t}\sum_{s=1}^{t-1} u_s \left(1 - \sum_{s=1}^{t-1} z_s u_s\right)$$

`// Predict and get feedback`

Predict with $Y_t = z_t \widetilde{Y}_t$

Pay $\ell_t\big([Y_t]_{i_t:}\big)$ and receives $g_t \in \partial \ell_t\big([Y_t]_{i_t:}\big)$

Set $u_t = \frac{\sqrt{N}}{\sqrt{2}L}\big\langle [\widetilde{Y}_t]_{i_t:}, g_t \big\rangle$

$$\le \sum_{j=1}^N \left[6 \max_{i\in\mathcal{N}_j} w_{ij} \left(\sqrt{1+\sigma_j^2(N_j-1)} + \ln N_j\right) \sqrt{\sum_{i\in\mathcal{N}_j} T_i} + 2N_j^2 \frac{\ln^2 T}{\varepsilon'} + 4d\sqrt{2N_j}\sqrt{1+\sigma_j^2(N_j-1)}\,\frac{\ln^2 T}{\varepsilon'} \right]$$

$$\stackrel{\widetilde{\mathcal{O}}}{=} \sum_{j=1}^N \max_{i\in\mathcal{N}_j} w_{ij} \sqrt{1+\sigma_j^2(N_j-1)} \sqrt{\sum_{i\in\mathcal{N}_j} T_i} + \frac{dN_{\max}^4}{\varepsilon} N \ln^2 T\,.$$

□

C Additional Results related to Section 5

In this section, we provide more details about the algorithm used to run Section 5's experiments. In Section 5, we use Krichevsky-Trofimov's Algorithm (hereafter abbreviated KT) instead of `Hedge` in order to obtain a σ-adaptive algorithm. Practically, this means that `AlgoClique` is set as the algorithm described in Algorithm 5. In the next theorem, we recall the arguments developed in Cesa-Bianchi et al. (2022) and Orabona (2019) to establish the regret bound of Algorithm 5.

Theorem 13. *Let G be any graph. The regret of* `MT-FTRL` *with KT adaptive learning rate and with* $\beta_{t-1} = \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sum_{s\le t-1}\mathbb{I}\{i_s \in \mathcal{N}_j\}}$ *satisfies for all $U \in \mathcal{U}$*

$$R_T(U) = \mathcal{O}\left(\sum_{j\in[N]} \sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\right),$$

where we neglect $\log N$ and $\log T$ factors.

Proof. By Lemma 2, we have

$$R_T(U) \le \sum_{j=1}^N R_T^{\text{clique-}j}\big(U^{(j)}\big)\,,$$

where $R_T^{\text{clique-}j}$ the regret suffered by `MT-FTRL` on the linear losses $\langle w_{i_t j}\, g_t, \cdot\rangle$ over the rounds $t \le T$ such that $i_t \in \mathcal{N}_j$. We proceed by bounding all these terms individually. Let us focus on $R_T^{\text{clique-}j}$. To bound this term, we build upon Orabona (2019, Theorem 9.9). The parameter-independent one-dimensional algorithm $\mathcal{A}_{1d}$ is set as the Krichevsky-Trofimov (KT) algorithm, see Orabona (2019, Algorithm 9.2), and used to learn $\|U^{(j)}\|_{A_j}$. It produces the sequence $(z_t^{(j)})_{t\ge1}$. Instead, the algorithm $\mathcal{A}_{\mathcal{B}}$ is set as `MT-FTRL` on the ball of radius 1 with respect to norm $\|\cdot\|_{A_j}$. It produces the sequence $(\widetilde{Y}_t^{(j)})_{t\geq 1}$. The prediction at time step t is $Y_t^{(j)} = z_t^{(j)}\widetilde{Y}_t^{(j)}$. We can decompose the regret of Algorithm 5 as follows

$$
\begin{aligned}
R_T^{\text{clique-}j}(U^{(j)}) &\leq \sum_{t=1}^{T}\sum_{i\in\mathcal{N}_j}\big\langle w_{ij}g_t, [Y_t^{(j)}]_{i_t:} - U_{i_t:}^{(j)}\big\rangle \mathbb{I}\{i_t = i\}\\
&= \sum_{t:\ i_t\in\mathcal{N}_j}\big\langle w_{ij}g_t, z_t^{(j)}[\widetilde{Y}_t^{(j)}]_{i_t:} - U_{i_t:}^{(j)}\big\rangle\\
&= \sum_{t:\ i_t\in\mathcal{N}_j}\big\langle G_t^w, z_t^{(j)}\widetilde{Y}_t^{(j)} - U^{(j)}\big\rangle\\
&= \sum_{t:\ i_t\in\mathcal{N}_j}\big\langle G_t^w, z_t^{(j)}\widetilde{Y}_t^{(j)} - \|U^{(j)}\|_{A_j}\,\widetilde{Y}_t^{(j)}\big\rangle + \sum_{t:\ i_t\in\mathcal{N}_j}\big\langle G_t^w, \|U^{(j)}\|_{A_j}\,\widetilde{Y}_t^{(j)} - U^{(j)}\big\rangle\\
&= \frac{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}{\sqrt{N_j}}\sum_{t:\ i_t\in\mathcal{N}_j}\frac{\sqrt{N_j}}{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}\big\langle G_t^w, \widetilde{Y}_t^{(j)}\big\rangle\big(z_t^{(j)} - \|U^{(j)}\|_{A_j}\big) \qquad (44)\\
&\quad + \|U^{(j)}\|_{A_j}\sum_{t:\ i_t\in\mathcal{N}_j}\left\langle G_t^w, \widetilde{Y}_t^{(j)} - \frac{U^{(j)}}{\|U^{(j)}\|_{A_j}}\right\rangle\\
&\leq \frac{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}{\sqrt{N_j}}\ \text{Regret}_T^{\mathcal{A}_{1d}}\big(\|U^{(j)}\|_{A_j}\big) + \|U^{(j)}\|_{A_j}\ \text{Regret}_T^{\mathcal{A}_{\mathcal{B}}}\left(\frac{U^{(j)}}{\|U^{(j)}\|_{A_j}}\right), \qquad (45)
\end{aligned}
$$

where G_t^w is the matrix containing only zeros except at row i_t, which is equal to $w_{i_tj}g_t$, and $\text{Regret}_T^{\mathcal{A}_{1d}}$ and $\text{Regret}_T^{\mathcal{A}_{\mathcal{B}}}$ denote upper bounds on the regrets (with linear losses) of algorithms $\mathcal{A}_{1d}$ and $\mathcal{A}_{\mathcal{B}}$ respectively. Note that the last inequality holds as we do have

$$
\left|\frac{\sqrt{N_j}}{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}\big\langle G_t^w, \widetilde{Y}_t^{(j)}\big\rangle\right| \leq \frac{\sqrt{N_j}}{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}\,\|G_t^w\|_{A_j^{-1}}\,\|\widetilde{Y}_t^{(j)}\|_{A_j} = \frac{w_{i_tj}\sqrt{N_j}}{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{2}}\sqrt{[A_j^{-1}]_{i_ti_t}}\,\|g_t\| \leq 1\,.
$$

Now, using Orabona (2019, Section 9.2.1), we have that there exists a universal constant C_0 such that

$$
\begin{aligned}
\text{Regret}_T^{\mathcal{A}_{1d}}\big(\|U^{(j)}\|_{A_j}\big) &\leq \|U^{(j)}\|_{A_j}\sqrt{4\sum_{i\in\mathcal{N}_j}T_i\ln\Big(1 + C_0\|U^{(j)}\|_{A_j}\sum_{i\in\mathcal{N}_j}T_i\Big)} + 1\\
&\leq \sqrt{N_j\big(1+\sigma^2(N_j-1)\big)}\sqrt{4\Big(\sum_{i\in\mathcal{N}_j}T_i\Big)\ln\Big(1 + C_0N_j\sum_{i\in\mathcal{N}_j}T_i\Big)} + 1\,, \qquad (46)
\end{aligned}
$$

where we used Cesa-Bianchi et al. (2022, Equation (21)) to compute $\|U^{(j)}\|_{A_j}$. On the other hand, `MT-FTRL` is an instance of `FTRL` with regularizer $X \mapsto (1/2)\|X\|_{A_j}^2$ and learning rates $\eta_{t-1} = 1/\beta_{t-1}$. Its regret on the unit ball with respect to $\|\cdot\|_{A_j}$ can thus be bounded (see e.g., Orabona 2019, Corollary 7.9) by

$$
\begin{aligned}
\text{Regret}_T^{\mathcal{A}_{\mathcal{B}}}\left(\frac{U^{(j)}}{\|U^{(j)}\|_{A_j}}\right) &\leq \frac{\Big\|\frac{U^{(j)}}{\|U^{(j)}\|_{A_j}}\Big\|_{A_j}}{2\,\eta_{T-1}} + \frac{1}{2}\sum_{t:\ i_t\in\mathcal{N}_j}\eta_{t-1}\big\|G_t^w\big\|_{A_j^{-1}}^2\\
&\leq \frac{1}{2\eta_{T-1}} + \frac{1}{2}\sum_{t:\ i_t\in\mathcal{N}_j}\eta_{t-1}[A_j^{-1}]_{i_ti_t}w_{i_tj}^2\|g_t\|_2^2\\
&\leq \frac{\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{N_j}}{2}\sqrt{\sum_{i\in\mathcal{N}_j}T_i} + \frac{1}{\max_{i\in\mathcal{N}_j} w_{ij}}\sum_{t:\ i_t\in\mathcal{N}_j}\frac{w_{i_tj}^2}{\sqrt{1+\sum_{s\leq t-1}\mathbb{I}\{i_s\in\mathcal{N}_j\}}} \qquad (47)\\
&\leq \sqrt{N_j}\max_{i\in\mathcal{N}_j} w_{ij}\frac{\sqrt{\sum_{i\in\mathcal{N}_j}T_i}}{2} + \frac{2\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{N_j}\sqrt{\sum_{i\in\mathcal{N}_j}T_i}}{N_j}\\
&= 2\max_{i\in\mathcal{N}_j} w_{ij}\sqrt{\frac{\sum_{i\in\mathcal{N}_j}T_i}{N_j}}\,, \qquad (48)
\end{aligned}
$$

where the third inequality comes from $[{A_j}^{-1}]_{ii} = \frac{2}{N_j+1}$ (see for example computations in Appendix A.2 of Cesa-Bianchi et al. (2022)). Substituting (46) and (48) into (45), we obtain

$$
\begin{aligned}
&R_T(U^{(j)}) \\
&\leq \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\left(2+\sqrt{8\ln\left(1+C_0 N_j\sum_{i\in\mathcal{N}_j} T_i\right)}\right) + \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i} \\
&\leq \max_{i\in\mathcal{N}_j} w_{ij}\sqrt{1+\sigma_j^2(N_j-1)}\sqrt{\sum_{i\in\mathcal{N}_j} T_i}\left(3+\sqrt{8\ln\left(1+C_0 N_j\sum_{i\in\mathcal{N}_j} T_i\right)}\right) .
\end{aligned}
$$

□

Remark 3. *Note that in the stochastic case, Equation (47) shows that we could use* $\beta_{t-1} = \sqrt{\sum_{i\in\mathcal{N}_j} \frac{q_i}{Q_j}\, w_{ij}^2}$ $\sqrt{1+\sum_{s\leq t-1}\mathbb{I}\left\{i_s\in\mathcal{N}_j\right\}}$ *for each agent* j *as in Theorem 5, therefore reducing the expectation of the second term of the regret of Equation (45) to a* $\widetilde{\mathcal{O}}\left(\sum_{j=1}^N \sqrt{\sum_{i\in\mathcal{N}_j} q_i\, w_{ij}^2}\sqrt{1+\sigma_j^2(N_j-1)}\right)\sqrt{T}$. *However, that would not change the first term in the regret of Equation (45), related to the regret of the KT-algorithm itself.*