References

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

Noga Alon, Nicolo Cesa-Bianchi, Claudio Gentile, Shie Mannor, Yishay Mansour, and Ohad Shamir. Non-stochastic multi-armed bandits with graph-structured feedback. *SIAM Journal on Computing*, 46(6): 1785–1826, 2017.

Nicolò Cesa-Bianchi, Tommaso Cesari, and Claire Monteleoni. Cooperative online learning: Keeping your neighbors updated. In *Algorithmic Learning Theory*, pages 234–250. PMLR, 2020.

Yu-Guan Hsieh, Franck Iutzeler, Jérôme Malick, and Panayotis Mertikopoulos. Multi-agent online optimization with delays: Asynchronicity, adaptivity, and optimism. *Journal of Machine Learning Research*, 23(78): 1–49, 2022.

Nicolò Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. Delay and cooperation in nonstochastic bandits. *Journal of Machine Learning Research*, 20(17):1–38, 2019.

Yogev Bar-On and Yishay Mansour. Individual regret in cooperative nonstochastic multi-armed bandits. *Advances in Neural Information Processing Systems*, 32, 2019a.

Nathan Linial. Locality in distributed graph algorithms. *SIAM Journal on computing*, 21(1):193–201, 1992.

Jukka Suomela. Survey of local algorithms. *ACM Computing Surveys (CSUR)*, 45(2):1–40, 2013.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

Andras Gyorgy and Pooria Joulani. Adapting to delays and data in adversarial multi-armed bandits. In *International Conference on Machine Learning*, pages 3988–3997. PMLR, 2021.

Mark Herbster, Stephen Pasteris, Fabio Vitale, and Massimiliano Pontil. A gang of adversarial bandits. *Advances in Neural Information Processing Systems*, 34, 2021.

Yogev Bar-On and Yishay Mansour. Individual regret in cooperative nonstochastic multi-armed bandits. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019b.

Shinji Ito, Daisuke Hatano, Hanna Sumita, Kei Takemura, Takuro Fukunaga, Naonori Kakimura, and Ken-Ichi Kawarabayashi. Delay and cooperation in nonstochastic linear bandits. *Advances in Neural Information Processing Systems*, 33:4872–4883, 2020.

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

Abhimanyu Dubey et al. Kernel methods for cooperative multi-agent contextual bandits. In *International Conference on Machine Learning*, pages 2740–2750. PMLR, 2020a.

Tomoki Nakamura, Naoki Hayashi, and Masahiro Inuiguchi. Cooperative learning for adversarial multi-armed bandit on open multi-agent systems. *IEEE Control Systems Letters*, 2023.

Ravi Kumar Kolla, Krishna Jagannathan, and Aditya Gopalan. Collaborative learning of stochastic bandits over a social network. *IEEE/ACM Transactions on Networking*, 26(4):1782–1795, 2018.

Yu-Zhen Janice Chen, Stephen Pasteris, Mohammad Hajiesmaili, John Lui, Don Towsley, et al. Cooperative stochastic bandits with asynchronous agents and constrained feedback. *Advances in Neural Information Processing Systems*, 34, 2021.

Lin Yang, Yu-zhen Janice Chen, Mohammad Hajiesmaili, John Lui, and Don Towsley. Distributed bandits with heterogeneous agents. *arXiv preprint arXiv:2201.09353*, 2022.

David Martínez-Rubio, Varun Kanade, and Patrick Rebeschini. Decentralized cooperative stochastic bandits. *Advances in Neural Information Processing Systems*, 32, 2019.

Udari Madhushani, Abhimanyu Dubey, Naomi Leonard, and Alex Pentland. One more step towards reality: Cooperative bandits with imperfect communication. *Advances in Neural Information Processing Systems*, 34, 2021.

Abhimanyu Dubey et al. Cooperative multi-agent bandits with heavy tails. In *International Conference on Machine Learning*, pages 2730–2739. PMLR, 2020b.

Ronshee Chawla, Abishek Sankararaman, Ayalvadi Ganesh, and Sanjay Shakkottai. The gossiping insert-eliminate algorithm for multi-agent bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 3471–3481. PMLR, 2020.

Abhimanyu Dubey and Alex Pentland. Private and byzantine-proof cooperative decision-making. In *AAMAS*, pages 357–365, 2020.

Anusha Lalitha and Andrea Goldsmith. Bayesian algorithms for decentralized stochastic bandits. *IEEE Journal on Selected Areas in Information Theory*, 2(2):564–583, 2021.

Jiafan He, Tianhao Wang, Yifei Min, and Quanquan Gu. A simple and provably efficient algorithm for asynchronous federated contextual linear bandits. *Advances in neural information processing systems*, 35: 4762–4775, 2022.

Chuanhao Li and Hongning Wang. Asynchronous upper confidence bound algorithms for federated linear bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 6529–6553. PMLR, 2022.

Lydia T Liu, Feng Ruan, Horia Mania, and Michael I Jordan. Bandit learning in decentralized matching markets. *The Journal of Machine Learning Research*, 22(1):9612–9645, 2021.

Pooria Joulani, András György, and Csaba Szepesvári. A modular analysis of adaptive (non-) convex optimization: Optimism, composite objectives, variance reduction, and variational bounds. *Theoretical Computer Science*, 808:108–138, 2020.

R. Stanton Hales. Numerical invariants and the strong product of graphs. *Journal of Combinatorial Theory, Series B*, 15(2):146–155, 1973.

Antonio Acín, Runyao Duan, David E Roberson, Ana Belén Sainz, and Andreas Winter. A new property of the Lovász number and duality relations between graph parameters. *Discrete Applied Mathematics*, 216: 489–501, 2017.

Moshe Rosenfeld. On a problem of C.E. Shannon in graph theory. *Proceedings of the American Mathematical Society*, 18(2):315–319, 1967.

Noga Alon, Nicolo Cesa-Bianchi, Ofer Dekel, and Tomer Koren. Online learning with feedback graphs: Beyond bandits. In *Conference on Learning Theory*, pages 23–35. PMLR, 2015.

Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and trends in Machine Learning*, 4(2):107–194, 2012.

Francesco Orabona, Koby Crammer, and Nicolò Cesa-Bianchi. A generalized online mirror descent with applications to classification and regression. *Machine Learning*, 99(3):411–435, 2015.

A Proof of Lemma 2

Lemma 2. *For any feedback graph F, for any $q > 0$, and for any $T \geq \max\{0.0064 \cdot \alpha(F)^3, \frac{1}{q^3}\}$, if each round $t \in [T]$ is independently skipped with probability q, then*

$$\inf_{\text{ALG}} \sup_{\ell} R_T^{\text{sa}}(q, \text{ALG}, \ell) \overset{\Omega}{=} \sqrt{\alpha(F) q T} \ .$$

Proof. Let $\mathcal{T}$ be the (random) set of times $\{t \in [T] \mid A_t = 1\}$ and let $\tau_1 < \tau_2 < \cdots < \tau_{|\mathcal{T}|}$the (random) elements of $\mathcal{T}$ in increasing order. Fix an online learning algorithm ALG and a sequence $\ell = \big(\ell_t\big)_{t\in[T]}$ of losses. For any random variable J (later, J and the corresponding "hard" instance will be those used in the lower bound for online learning with feedback graphs: (Alon et al., 2017, Theorem 5)). Then

$$\begin{aligned} R_T(q, \text{ALG}, \ell) &= \max_{i\in[K]} \mathbb{E}\left[\sum_{t=1}^{T} \big(\ell_t(I_{N_t+1}) - \ell_t(i)\big)\mathbb{I}\{A_t = 1\}\right] = \max_{i\in[K]} \mathbb{E}\left[\sum_{s\in[|\mathcal{T}|]} \big(\ell_{\tau_s}(I_s) - \ell_{\tau_s}(i)\big)\right] \\ &\geq \mathbb{E}\left[\sum_{s\in[|\mathcal{T}|]} \big(\ell_{\tau_s}(I_s) - \ell_{\tau_s}(J)\big)\right] \\ &= \sum_{n\in[T]} \sum_{\substack{\mathcal{T}_0\subset[T]\\ |\mathcal{T}_0|=n}} \mathbb{E}\left[\sum_{s\in[|\mathcal{T}|]} \big(\ell_{\tau_s}(I_s) - \ell_{\tau_s}(J)\big) \mid \mathcal{T} = \mathcal{T}_0, |\mathcal{T}_0| = n\right] \mathbb{P}\left(\mathcal{T} = \mathcal{T}_0, |\mathcal{T}_0| = n\right) \end{aligned}$$

Then, we recognize that the conditional expectation in the previous formula is the expected regret for single-agent online learning with feedback graph. Therefore, from (Alon et al., 2017, Theorem 5) we get that, letting $C_1 = (8/100)^2$ and , for all $T \geq C_1 \alpha(F)^3$,

$$\begin{aligned} \inf_{\text{ALG}} \sup_{\ell} R_T(q, \text{ALG}, \ell) &\geq \sum_{n\in[T]} \sum_{\substack{\mathcal{T}_0\subset[T]\\ |\mathcal{T}_0|=n}} \left(\varepsilon n \left(\frac{1}{2} - 2\varepsilon\sqrt{\frac{n}{\alpha(F)}}\right)\right) \mathbb{P}\left(\mathcal{T} = \mathcal{T}_0, |\mathcal{T}_0| = n\right) \\ &= \sum_{n\in[T]} \left(\varepsilon n \left(\frac{1}{2} - 2\varepsilon\sqrt{\frac{n}{\alpha(F)}}\right)\right) \mathbb{P}\left(|\mathcal{T}| = n\right) \\ &= \sum_{n\in[T]} \varepsilon n \left(\frac{1}{2} - 2\varepsilon\sqrt{\frac{n}{\alpha(F)}}\right) f_{\text{Bin}(q,T)}(n) \end{aligned}$$

where $f_{\text{Bin}(q,T)}$ is the p.m.f. of a Binomial random variable with parameters p, T. We want $\varepsilon = \varepsilon^*(p, T)$ that maximizes that expression. Therefore, by defining $g(\varepsilon) = a\varepsilon + b\varepsilon^2$ as the following quadratic polynomial in ε

$$\begin{aligned} g(\varepsilon) &= \sum_{n\in[T]} \left(\varepsilon n \left(\frac{1}{2} - 2\varepsilon\sqrt{\frac{n}{\alpha(F)}}\right)\right) f_{\text{Bin}(q,T)}(n) \\ &= \sum_{n\in[T]} \left(\frac{n}{2} f_{\text{Bin}(q,T)}(n)\right)\varepsilon - 2\sum_{m\in[T]} \left(\frac{m^{3/2}}{\alpha(F)^{1/2}} f_{\text{Bin}(q,T)}(m)\right)\varepsilon^2 \end{aligned}$$

where $a = \sum_{n\in[T]} \left(\frac{n}{2} f_{\text{Bin}(q,T)}(n)\right)$ and $b = -2\sum_{m\in[T]} \left(\frac{m^{3/2}}{\alpha(F)^{1/2}} f_{\text{Bin}(q,T)}(m)\right)$. We find that the maximum of g is achieved in $-\frac{a}{2b}$, which gives the optimal value

$$\varepsilon^*(p, T) = \frac{\sum_{n\in[T]} \left(\frac{n}{2} f_{\text{Bin}(q,T)}(n)\right)}{4\sum_{m\in[T]} \left(\frac{m^{3/2}}{\alpha(F)^{1/2}} f_{\text{Bin}(q,T)}(m)\right)}$$

and the function g evaluated at the optimal value is equal to $g^* = -\frac{a^2}{4b}$, i.e.,

$$g^* = g(\varepsilon^*(p,T)) = g\left(\frac{\sum_{n\in[T]}\left(\frac{n}{2}f_{\mathrm{Bin}(q,T)}(n)\right)}{4\sum_{m\in[T]}\left(\frac{m^{3/2}}{\alpha(F)^{1/2}}f_{\mathrm{Bin}(q,T)}(m)\right)}\right) = \frac{\sqrt{\alpha(F)}}{8}\frac{\left(\sum_{n\in[T]}nf_{\mathrm{Bin}(q,T)}(n)\right)^2}{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}$$

Therefore, we can lower bound the regret with g^* and obtain

$$\begin{aligned}\inf_{\textsc{Alg}}\sup_{\ell} R_T(q,\textsc{Alg},\ell) &\geq \frac{\sqrt{\alpha(F)}}{8}\frac{\left(\sum_{n\in[T]}nf_{\mathrm{Bin}(q,T)}(n)\right)^2}{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)} = \frac{\sqrt{\alpha(F)qT}}{8}\frac{(qT)^2}{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}\\ &= \frac{\sqrt{\alpha(F)qT}}{8}\frac{(qT)^{3/2}}{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}\end{aligned}$$

where in the first equality we substituted the expected value of a binomial distribution of parameter q.

We now want to prove the existence of a constant $c > 0$ such that, for every $q > 0$ and every $T \geq \frac{1}{q^3}$

$$\frac{(qT)^{3/2}}{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)} \geq c\ , \quad \text{or equivalently} \quad \frac{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}} \leq c$$

We split the sum over $m \in [T]$ into two blocks, the first for $1 \leq m \leq c_2\lfloor qT\rfloor$ and the second for $c_2\lfloor qT\rfloor < m \leq T$ for a constant $c_2 = \left\lceil \frac{qT}{\lfloor qT\rfloor}\left(\frac{1}{q}\left(\sqrt{\frac{1}{2T}\ln\left(\frac{T^{3/2}}{c_1}\right)} - \frac{1}{T}\right) + 1\right)\right\rceil$ and with $c_1 > 0$:

$$\frac{\sum_{m\in[T]}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}} = \frac{\sum_{m=1}^{c_2\lfloor qT\rfloor}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}} + \frac{\sum_{m>c_2\lfloor qT\rfloor}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}}\ . \tag{9}$$

The idea is to choose the split point $c_2\lfloor qT\rfloor$ so that we can upper bound the tail mass using Hoeffding's inequality. Hoeffding's inequality yields the simple bound $F_{\mathrm{Bin}(q,T)}(m) \leq \exp\left(-2T\left(q - \frac{m}{T}\right)^2\right)$, and together with symmetry properties of the binomial distribution $1 - F_{\mathrm{Bin}(q,T)}(m) = F_{\mathrm{Bin}(1-q,T)}(T-m)$ we obtain a bound on the upper tail. This contribution compensates exactly the term $q^{3/2}$ left at the denominator for $T \geq 1/q$, leaving just the constant c_1:

$$\begin{aligned}\frac{\sum_{m>c_2\lfloor qT\rfloor}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}} &\leq \frac{e^{-2T\left((1-q)-\frac{T-(c_2\lfloor qT\rfloor+1)}{T}\right)^2}T^{3/2}}{(qT)^{3/2}}\\ &= \frac{e^{-2T\left(q\left(c_2\frac{\lfloor qT\rfloor}{qT}-1\right)+\frac{1}{T}\right)^2}}{q^{3/2}}\\ &= \frac{c_1}{(qT)^{3/2}}\\ &\leq c_1\ .\end{aligned}$$

For the first term in Equation (9), we upper bound the lower tail simply by one:

$$\sum_{m=1}^{c_2\lfloor qT\rfloor}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right) \leq c_2\lfloor qT\rfloor \cdot F_{\mathrm{Bin}(q,T)}(m) \leq c_2\lfloor qT\rfloor$$

We conclude by proving that the first term in Equation (9) is bounded by a constant. If we take $m \leq c_2\lfloor qT\rfloor$ we obtain for $T \geq 2/q$ and $c_1 \geq 1$

$$\frac{\sum_{m=1}^{c_2\lfloor qT\rfloor}\left(m^{3/2}f_{\mathrm{Bin}(q,T)}(m)\right)}{(qT)^{3/2}} \leq \frac{(c_2\lfloor qT\rfloor)^{3/2}}{(qT)^{3/2}}$$

$$
\begin{aligned}
&\le (c_2)^{3/2} \\
&\le \left(1 + \frac{qT}{\lfloor qT \rfloor}\left(\frac{1}{q}\left(\sqrt{\frac{1}{2T}\ln\left(\frac{T^{3/2}}{c_1}\right)} - \frac{1}{T}\right) + 1\right)\right)^{3/2} \\
&\le \left(2 + \frac{3\sqrt{3}}{8}\frac{1}{q\sqrt{T}}\left(\sqrt{\log(T)}\right)\right)^{3/2} \\
&\le \left(2 + \frac{3\sqrt{3}}{8}\right)^{3/2}\left(\frac{1}{q\sqrt{T}}\sqrt{\ln\left(\frac{T^{3/2}}{c_1}\right)}\right)^{3/2} \\
&\le 4.32\left(\frac{1}{q\sqrt{T}}\sqrt{\ln\left(\frac{T^{3/2}}{c_1}\right)}\right)^{3/2} \\
&\le 4.32\,(\ln T)^{3/4}\,\frac{1}{(q^2T)^{3/4}} \\
&\le 4.32\frac{(\ln T)^{3/4}}{T^{1/4}} \\
&\le 4.32 \cdot 1.08 \le 5
\end{aligned}
$$

where in the third-last inequality we used $q \ge \frac{1}{T^{1/3}}$. Putting everything together and letting $c_1 = 1$ yields

$$\inf_{\textsc{Alg}} \sup_{\ell} R_T(q, \textsc{Alg}, \ell) \ge \frac{3}{4}\sqrt{\alpha(F)qT}$$

□

B Graph-theoretic results

In this section, we present a general version of a graph-theoretic lemma (Lemma 6) that is crucial for our positive results in Section 4. Before stating it, we recall a few known results.

The first result is a direct consequence of (Alon et al., 2017, Lemma 10) specialized to undirected graphs.

Lemma 3. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph containing all self-loops and $\alpha_d(\mathcal{G})$ its d-th independence number. For all $i \in \mathcal{V}$, let $\mathcal{N}_d^{\mathcal{G}}(i)$ be the d-th neighborhood of i, $p(i) \ge 0$, and $P(i) = \sum_{j \in \mathcal{N}_d^{\mathcal{G}}(i)} p(j) > 0$. Then*

$$\sum_{i \in \mathcal{V}} \frac{p(i)}{P(i)} \le \alpha_d(\mathcal{G})$$

Proof. Initialize $V_1 = \mathcal{V}$, fix $j_1 \in \operatorname{argmin}_{j \in V_1} P(j)$, and denote $V_2 = \mathcal{V} \setminus \mathcal{N}(j_1)$. For $k \ge 2$ fix $j_k \in \operatorname{argmin}_{j \in V_k} P(j)$ and shrink $V_{k+1} = V_k \setminus \mathcal{N}(j_k)$ until $V_{k+1} = \varnothing$. Since $\mathcal{G}$ is undirected $j_k \notin \bigcup_{s=1}^{k-1} \mathcal{N}(j_s)$, therefore the number m of times that an action can be picked this way is upper bounded by α. Denoting $\mathcal{N}'(j_k) = V_k \cap \mathcal{N}(j_k)$ this implies

$$\sum_{i \in \mathcal{V}} \frac{p(i)}{P(i)} = \sum_{k=1}^{m} \sum_{i \in \mathcal{N}'(j_k)} \frac{p(i)}{P(i)} \le \sum_{k=1}^{m} \sum_{i \in \mathcal{N}'(j_k)} \frac{p(i)}{P(j_k)} \le \sum_{k=1}^{m} \frac{\sum_{i \in \mathcal{N}(j_k)} p(i)}{P(j_k)} = m \le \alpha$$

concluding the proof. □

The following result, known as the inequality of arithmetic and geometric means, or simply AM-GM inequality, is used in the proofs of Lemmas 1, 5, and 6.

Lemma 4 (AM-GM inequality)**.** *For any $x_1, \dots, x_r \in [0, \infty)$,*

$$\frac{x_1 + \cdots + x_r}{r} \ge (x_1 \cdot \cdots \cdot x_r)^{1/r}$$

Proof. By Jensen's inequality,

$$\ln\left(\frac{1}{r}\sum_{i=1}^{r} x_i\right) \geq \sum_{i=1}^{r}\frac{1}{r}\ln\left(x_i\right) = \sum_{i=1}^{r}\ln\left(x_i^{1/r}\right) = \ln\left(\prod_{i=1}^{r} x_i^{1/r}\right)$$

□

The second result is proven in (Cesa-Bianchi et al., 2019, Lemma 3), but here we give a slightly different proof based on the AM-GM inequality.

Lemma 5. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph containing all self-loops and $\alpha_d(\mathcal{G})$ its d-th independence number. For all $v \in \mathcal{V}$, let $\mathcal{N}_d^{\mathcal{G}}(v)$ be the d-th neighborhood of v, $c(v) \geq 0$, and $C(v) = 1 - \prod_{w \in \mathcal{N}_d^{\mathcal{G}}(v)}\big(1 - c(w)\big) > 0$. Then*

$$\sum_{v \in \mathcal{V}} \frac{c(v)}{C(v)} \leq \frac{1}{1 - e^{-1}}\left(\alpha_d(\mathcal{G}) + \sum_{v \in \mathcal{V}} c(v)\right)$$

Proof. Set for brevity $P(v) = \sum_{w \in \mathcal{N}_d^{\mathcal{G}}(v)} c(w)$. Then we can write

$$\sum_{v \in \mathcal{V}} \frac{c(v)}{C(v)} = \underbrace{\sum_{v \in \mathcal{V}\,:\,P(v) \geq 1} \frac{c(v)}{C(v)}}_{\text{(I)}} + \underbrace{\sum_{v \in \mathcal{V}\,:\,P(v) < 1} \frac{c(v)}{C(v)}}_{\text{(II)}}$$

and proceed by upper bounding the two terms (I) and (II) separately. Let $r(v)$ be the cardinality of $\mathcal{N}_d^{\mathcal{G}}(v)$. We have, for any given $v \in \mathcal{V}$,

$$\begin{aligned}
\min\left\{C(v) \,:\, \sum_{w \in \mathcal{N}(v)} c(w) \geq 1\right\} &= \min\left\{C(v) \,:\, \sum_{w \in \mathcal{N}(v)} c(w) = 1\right\} \\
&= 1 - \max\left\{\prod_{w \in \mathcal{N}_d^{\mathcal{G}}(v)}\big(1 - c(w)\big) \,:\, \sum_{w \in \mathcal{N}(v)}\big(1 - c(w)\big) = r(v) - 1\right\} \\
&\geq 1 - \left(1 - \frac{1}{r(v)}\right)^{r(v)} \geq 1 - e^{-1}
\end{aligned}$$

where the first equality follows from the definition of $C(v)$ and the monotonicity of $x \mapsto 1 - x$, the first inequality is implied by the AM-GM inequality (Lemma 4), and the last one comes from $r(v) \geq 1$ (for $v \in \mathcal{N}_d^{\mathcal{G}}(v)$). Hence

$$\text{(I)} \leq \sum_{v \in \mathcal{V}\,:\,P(v) \geq 1} \frac{c(v)}{1 - e^{-1}} \leq \sum_{v \in \mathcal{V}} \frac{c(v)}{1 - e^{-1}}$$

As for (II), using the inequality $1 - x \leq e^{-x}$, $x \in \mathbb{R}$, with $x = c(w)$, we can write

$$C(v) \geq 1 - \exp\left(-\sum_{w \in \mathcal{N}_d^{\mathcal{G}}(v)} c(w)\right) = 1 - \exp\left(-P(v)\right)$$

Now, since in (II) we are only summing over v such that $P(v) < 1$, we can use the inequality $1 - e^{-x} \geq \left(1 - e^{-1}\right)x$, holding when $x \in [0, 1]$, with $x = P(v)$, thereby concluding that

$$C(v) \geq (1 - e^{-1})P(v)$$

Thus

$$\text{(II)} \leq \sum_{v \in \mathcal{V}\,:\,P(v) < 1} \frac{c(v)}{(1 - e^{-1})P(v)} \leq \frac{1}{1 - e^{-1}}\sum_{v \in \mathcal{V}} \frac{c(v)}{P(v)} \leq \frac{\alpha}{1 - e^{-1}}$$

where in the last step we used Lemma 3. □

We can now state a more general version of our key graph-theoretic result, which can be proved similarly to Lemma 1.

Lemma 6. *Let $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}_2)$ be two undirected graphs containing all self-loops and $\alpha(\mathcal{G}_1 \boxtimes \mathcal{G}_2)$ the independence number of their strong product $\mathcal{G}_1 \boxtimes \mathcal{G}_2$. For all $(i,j) \in \mathcal{V}_1 \times \mathcal{V}_2$, let also $\mathcal{N}_1^{\mathcal{G}_1}(i)$, $\mathcal{N}_1^{\mathcal{G}_2}(v)$, and $\mathcal{N}_1^{\mathcal{G}_1 \boxtimes \mathcal{G}_2}(i,v)$ be the first neighborhoods of i (in $\mathcal{G}_1$), v (in $\mathcal{G}_2$), and (i,v) (in $\mathcal{G}_1 \boxtimes \mathcal{G}_2$). If $\boldsymbol{w} = \big(w(j,u)\big)_{(j,u)\in\mathcal{V}_1\times\mathcal{V}_2}$ is an arbitrary matrix with non-negative entries such that $1 - \sum_{j\in\mathcal{N}_1^{\mathcal{G}_1}(i)} w(j,u) \geq 0$ for all $(i,u) \in \mathcal{V}_1 \times \mathcal{V}_2$ and $1 - \prod_{u\in\mathcal{N}_1^{\mathcal{G}_2}(v)} \big(1 - \sum_{j\in\mathcal{N}_1^{\mathcal{G}_1}(i)} w(j,u)\big) > 0$ for all $(i,v) \in \mathcal{V}_1 \times \mathcal{V}_2$, then*

$$\sum_{i\in\mathcal{V}_1}\sum_{v\in\mathcal{V}_2} \frac{w(i,v)}{1 - \prod_{u\in\mathcal{N}_1^{\mathcal{G}_2}(v)}\big(1 - \sum_{j\in\mathcal{N}_1^{\mathcal{G}_1}(i)} w(j,u)\big)} \leq \frac{e}{e-1}\left(\alpha(\mathcal{G}_1 \boxtimes \mathcal{G}_2) + \sum_{i\in\mathcal{V}_1}\sum_{v\in\mathcal{V}_2} w(i,v)\right)$$

B.1 Further discussion on $\mathscr{G}$

In general, $\alpha(N)\alpha(F) \leq \alpha\big(N \boxtimes F\big)$ holds for any arbitrary pairs of graphs N, F. Indeed, the Cartesian product $I \times J$ of an independent set I of N and an independent set J of F is an independent set of $N \boxtimes F$. There exist graphs N, F with $\alpha(N)\alpha(F) \ll \alpha\big(N \boxtimes F\big)$, but these appear to be quite rare and pathological cases. For the sake of completeness, we add an example of such a construction below. This shows that not all pairs of graphs belong to $\mathscr{G}$.

Example 1. *Take as the first graph $G_1 = (V_1, E_1)$, the cycle C_5 over 5 vertices. Then, for any $k \geq 2$, build $G_k = (V_k, E_k)$ inductively by replacing each vertex $v \in V_{k-1}$ by a copy of C_5 and each edge $e \in E_{k-1}$ by a copy of $K_{5,5}$ (the complete bipartite graph with partitions of size 5 and 5) between the two copies of C_5 that replaced its endpoints. It can be shown that $\alpha(G_k) = 2^k$ but $\alpha\big(G_k \boxtimes G_k\big) \geq 5^k \gg 4^k = \alpha(G_k)^2$.*

To see why, note first that $\alpha(C_5) = 2$ but $\alpha\big(C_5 \boxtimes C_5\big) \geq 5$, by choosing the independent set containing the 5 vertices $(1,1)$, $(2,3)$, $(3,5)$, $(4,2)$, $(5,4)$. For $k \geq 2$, $\alpha(G_k) = 2^k$ but we can take the analogous in $G_k \boxtimes G_k$ of the above independent set in $C_5 \boxtimes C_5$. This gives 5 sets S_1, S_2, S_3, S_4, S_5 of 25^{k-1} vertices each, with no edges between S_i, S_j when $i \neq j$. The subgraph of $G_k \boxtimes G_k$ induced by each S_i is simply the previous iteration G_{k-1} of this construction, and proceeding by induction we can find an independent subset of each S_i with 5^{k-1} vertices, giving a total of 5^k independent vertices.

C The upper bound of Cesa-Bianchi et al. (2020) for experts

(Cesa-Bianchi et al., 2020, Theorem 10) gives theoretical guarantees for the *average* regret over active agents. In this section, we briefly discuss how to convert their statement to a corresponding result for the *total* regret over active agents that is the focus of our present work.

Before stating the theorem, we recall that the *convex conjugate* $f^* \colon \mathbb{R}^d \to \mathbb{R}$ of a convex function $f \colon \mathbb{X} \to \mathbb{R}$ is defined, for any $\boldsymbol{x} \in \mathbb{R}^d$, by $f^*(\boldsymbol{x}) = \sup_{\boldsymbol{w}\in\mathbb{X}} \big(\boldsymbol{x}\cdot\boldsymbol{w} - f(\boldsymbol{w})\big)$. Moreover, given $\sigma > 0$, we say that f is *σ-strongly convex* on $\mathbb{X}$ with respect to a norm $\|\cdot\|$ if, for all $\boldsymbol{u}, \boldsymbol{w} \in \mathbb{X}$, we have $f(\boldsymbol{u}) \geq f(\boldsymbol{w}) + \nabla f(\boldsymbol{w})\cdot(\boldsymbol{u}-\boldsymbol{w}) + \frac{\sigma}{2}\|\boldsymbol{u}-\boldsymbol{w}\|^2$. The following well-known result can be found in (Shalev-Shwartz et al., 2012, Lemma 2.19 and subsequent paragraph).

Lemma 7. *Let $f \colon \mathbb{X} \to \mathbb{R}$ be a strongly convex function on $\mathbb{X}$. Then the convex conjugate f^* is everywhere differentiable on $\mathbb{R}^d$.*

The following result—see, e.g., (Orabona et al., 2015, bound (6) in Corollary 1 with F set to zero)—shows an upper bound on the regret of Algorithm 2 for single-agent online convex optimization with expert feedback.

Theorem 3. *Let $g \colon \mathbb{X} \to \mathbb{R}$ be a differentiable function σ-strongly convex with respect to $\|\cdot\|$. Then the regret of Algorithm 2 run with $g_t = \frac{\sqrt{t}}{\eta} g$, for $\eta > 0$, satisfies*

$$\sum_{t=1}^T \ell_t(\boldsymbol{x}_t) - \inf_{\boldsymbol{x}\in\mathbb{X}} \sum_{t=1}^T \ell_t(\boldsymbol{x}) \leq \frac{D}{\eta}\sqrt{T} + \frac{\eta}{2\sigma}\sum_{t=1}^T \frac{1}{\sqrt{t}} \|\nabla \ell_t\|_*^2$$

Algorithm 2:

input: σ_t-strongly convex regularizers $g_t \colon \mathbb{X} \to \mathbb{R}$ for $t = 1, 2, \dots$
initialization: $\boldsymbol{\theta}_1 = \mathbf{0} \in \mathbb{R}^d$
for $t = 1, 2, \dots$ **do**
  choose $\boldsymbol{w}_t = \nabla g_t^*(\boldsymbol{\theta}_t)$
  observe $\nabla \ell_t(\boldsymbol{w}_t) \in \mathbb{R}^d$
  update $\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \nabla \ell_t(\boldsymbol{w}_t)$

where $D = \sup g$ and $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. If $\sup \|\nabla \ell_t\|_* \le L$, then choosing $\eta = \sqrt{2\sigma D}/L$ gives $R_T \le L\sqrt{2DT/\sigma}$.

We can now present the equivalent of (Cesa-Bianchi et al., 2020, Theorem 10) for cooperative online covenx optimization with expert feedback (i.e., F is a clique) where $n = 1$ but the feedback is broadcast to first neighbor immediately after an action is played (rather than the following round).

Theorem 4. *Consider a network $N = (A, E_N)$ of agents. If all agents v run Algorithm 2 with an oblivious network interface and $g_t = \frac{\sqrt{t}}{\eta} g$, where $\|g_t\|_*$ is upper bounded by a constant $L > 0$, $\eta > 0$ is a learning rate, and the regularizer $g \colon \mathbb{X} \to \mathbb{R}$ is differentiable, σ-strongly convex with respect to some norm $\|\cdot\|$, and upper bounded by a constant M^2, then the network regret satisfies*

$$R_T \le \left(\frac{M^2}{\eta} + \frac{\eta L^2}{2\sigma} \right) \sqrt{2Q(\alpha(N) + Q)T}$$

For $\eta = \sqrt{2\sigma} M/L$, we have

$$R_T \le \left(\sqrt{2\sigma} LM\right) \sqrt{Q(\alpha(N) + Q)T}$$

Proof sketch. For any $\boldsymbol{x} \in \mathbb{X}$, agent v, and time t, let $\boldsymbol{x}_t(v)$ be the prediction made by v at time t, $r_t(v, \boldsymbol{x}) = \ell_t\big(\boldsymbol{x}_t(v)\big) - \ell_t(\boldsymbol{x})$, $Q_v = \Pr\big(v \in \bigcup_{w \in \mathcal{A}_t} \mathcal{N}_1^N(w)\big) = 1 - \prod_{w \in \mathcal{N}_1^N(v)} \big(1 - q(w)\big)$, and $A' := \big\{w \in A : q(w) > 0\big\}$. Proceeding as in (Cesa-Bianchi et al., 2020, Theorem 2) yields, for each $v \in A'$ and $\boldsymbol{x} \in \mathbb{X}$,

$$\mathbb{E}\left[\sum_{t=1}^{T} r_t(v, \boldsymbol{x})\right] \le \left(\frac{M^2}{\eta} + \frac{\eta L^2}{2}\right) \sqrt{\frac{T}{Q_v}} \tag{10}$$

Now, by the independence of the activations of the agents at time t and $\big(r_t(v, \boldsymbol{x})\big)_{v \in A', \boldsymbol{x} \in \mathbb{X}}$, we get

$$R_T = \sup_{\boldsymbol{x} \in \mathbb{X}} \sum_{v \in V'} q(v) \sum_{t=1}^{T} \mathbb{E}\big[r_t(v, \boldsymbol{x})\big] \tag{11}$$

Putting Equations (10) and (11) together and applying Jensen's inequality yields

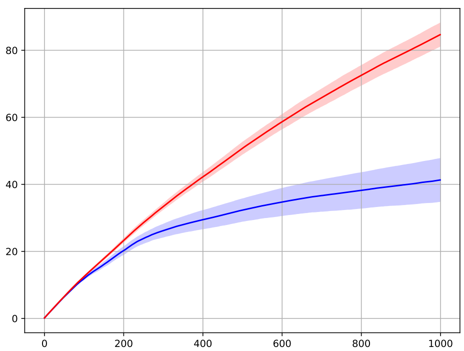
$$R_T \le \left(\sum_{v \in V'} q(v) \sqrt{\frac{1}{Q_v}}\right) \left(\frac{M^2}{\eta} + \frac{\eta L^2}{2}\right) \sqrt{T} \le \sqrt{Q \sum_{v \in V'} \frac{q_v}{Q_v}} \left(\frac{M^2}{\eta} + \frac{\eta L^2}{2}\right) \sqrt{T}$$

The proof is concluded by invoking Lemma 5. □

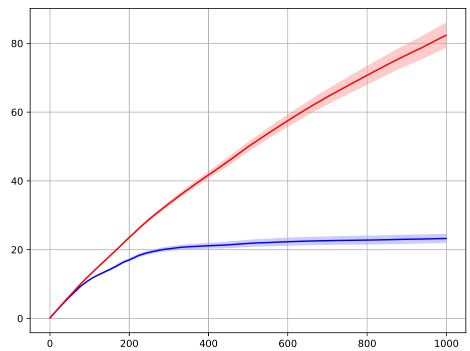
D Learning curves

We also plot the average regret R_T/Q against the number T of rounds. Our algorithm is the blue curve and the baseline is the red curve. Recall that these curves are averages over 20 repetitions of the same experiment (the shaded areas correspond to one standard deviation) where the stochasticity is due to the internal randomization of the algorithms. Experiments are designed to show the difference in performance when we allow agents to communicate and when we do not. The strong product captures in a mathematical form this difference in the regret bound for our algorithm, while the experiments here show it empirically. In particular, the bound for the case of no communication is bigger, and performances are worse in our simulations, as expected from theory.

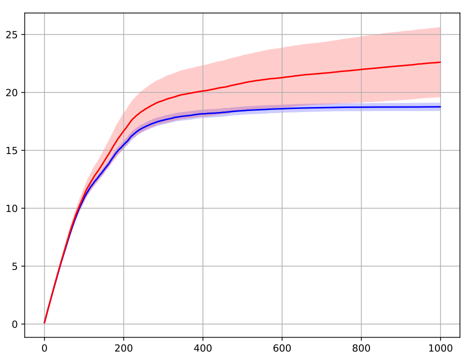
Experiments were run on a local cluster of CPUs (Intel Xeon E5-2623 v3, 3.00GHz), parallelizing the code over four cores. The run took approximately two hours.



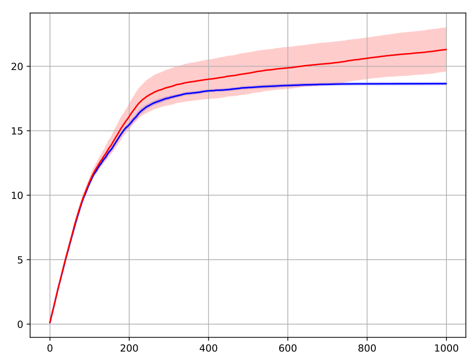
(a) $p_{ER}^N = 0.2$, $p_{ER}^F = 0.2$



(b) $p_{ER}^N = 0.8$, $p_{ER}^F = 0.2$

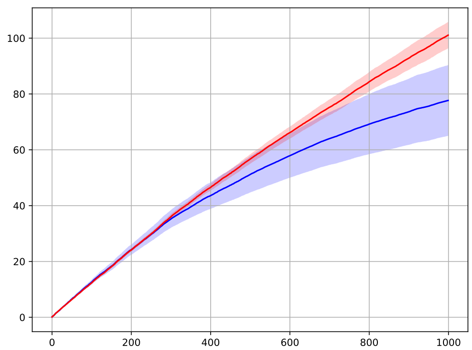


(c) $p_{ER}^N = 0.2$, $p_{ER}^F = 0.8$

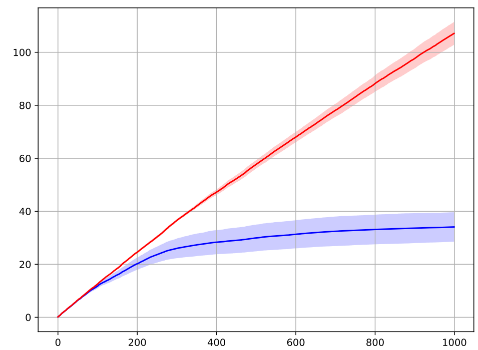


(d) $p_{ER}^N = 0.8$, $p_{ER}^F = 0.8$

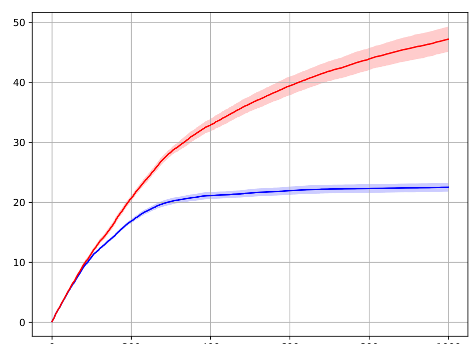
Figure 3: Average regret R_T/Q against $T = 1000$ of rounds. Activation probability $q = 1$.



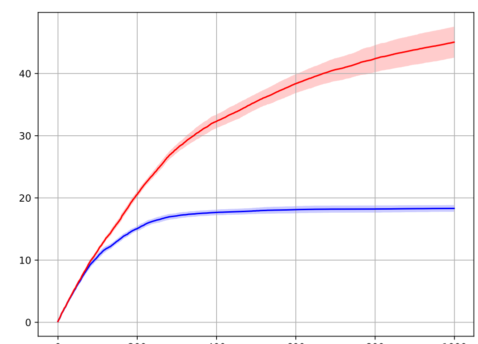
(a) $p^N_{ER} = 0.2$, $p^F_{ER} = 0.2$



(b) $p^N_{ER} = 0.8$, $p^F_{ER} = 0.2$

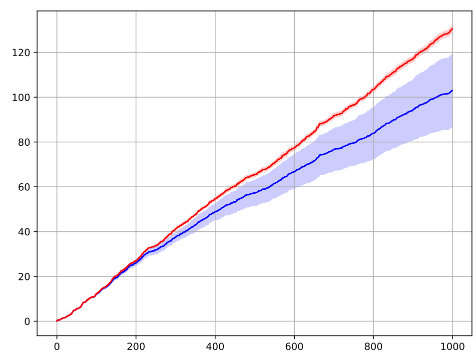


(c) $p^N_{ER} = 0.2$, $p^F_{ER} = 0.8$

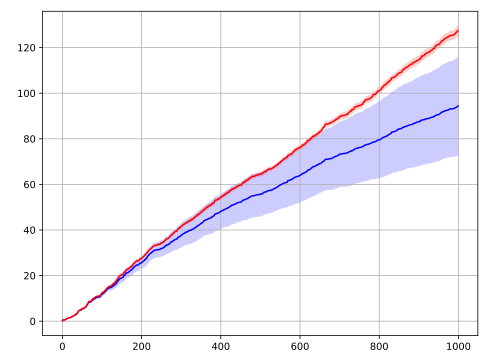


(d) $p^N_{ER} = 0.8$, $p^F_{ER} = 0.8$

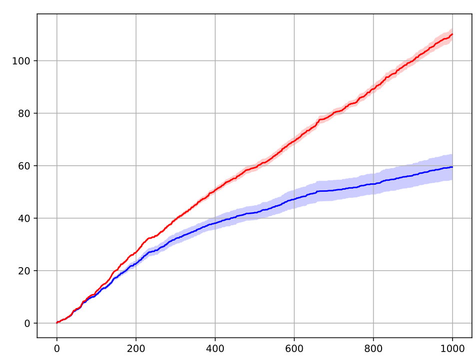
Figure 4: Average regret R_T/Q against $T = 1000$ of rounds. Activation probability $q = 0.5$.



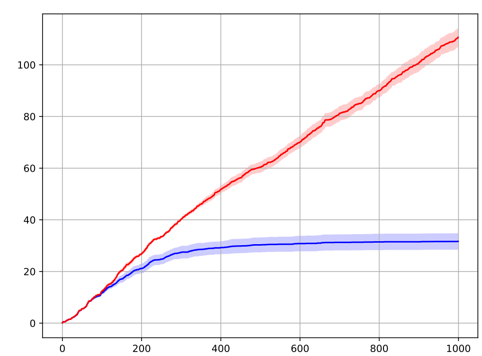
(a) $p_{ER}^N = 0.2$, $p_{ER}^F = 0.2$



(b) $p_{ER}^N = 0.8$, $p_{ER}^F = 0.2$



(c) $p_{ER}^N = 0.2$, $p_{ER}^F = 0.8$



(d) $p_{ER}^N = 0.8$, $p_{ER}^F = 0.8$

Figure 5: Average regret R_T/Q against $T = 1000$ of rounds. Activation probability $q = 0.05$.