Cooperation in Networks of Learning Agents

Nicolò Cesa-Bianchi

Università degli Studi di Milano

Tom Cesari (Univ. of Milano), Riccardo Della Vecchia (Univ. of Lille), Pierre Laforgue (Univ. of Milano), Andrea Paudice (Univ. of Milano and IIT), Massimiliano Pontil (IIT)

Themes of this talk

- Sequential decision-making: markets, sensors, user interactions

Themes of this talk

- Sequential decision-making: markets, sensors, user interactions
- Distributed learning systems (finance, recommendation, advertising, monitoring)

Themes of this talk

- Sequential decision-making: markets, sensors, user interactions
- Distributed learning systems (finance, recommendation, advertising, monitoring)
- The extent to which the amount of information received from the environment and from other agents affects how fast each agent can learn

Themes of this talk

- Sequential decision-making: markets, sensors, user interactions
- Distributed learning systems (finance, recommendation, advertising, monitoring)
- The extent to which the amount of information received from the environment and from other agents affects how fast each agent can learn
- We study this problem in an abstract graph-theoretic online learning framework

The nonstochastic bandit problem

- K actions
- Unknown deterministic assignment of losses to actions $\boldsymbol{\ell}_t = \big(\ell_t(1), \dots, \ell_t(K)\big) \in [0,1]^K$ for each time step t

? ? ? ? ? ? ? ? ? ?

For $t = 1, 2, \dots$

The nonstochastic bandit problem

- K actions
- Unknown deterministic assignment of losses to actions $\boldsymbol{\ell}_t = (\ell_t(1), \dots, \ell_t(K)) \in [0,1]^K$ for each time step t

For $t = 1, 2, \dots$

1. Player picks an action I_t (possibly using randomization) and incurs loss $\ell_t(I_t)$

The nonstochastic bandit problem

- K actions
- Unknown deterministic assignment of losses to actions $\boldsymbol{\ell}_t = (\ell_t(1), \dots, \ell_t(K)) \in [0,1]^K$ for each time step t

For $t = 1, 2, \dots$

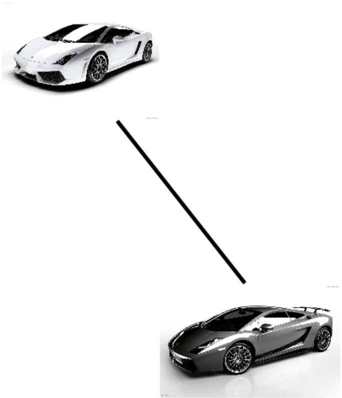
1. Player picks an action I_t (possibly using randomization) and incurs loss $\ell_t(I_t)$
2. Feedback from environment: player observes $\ell_t(I_t)$

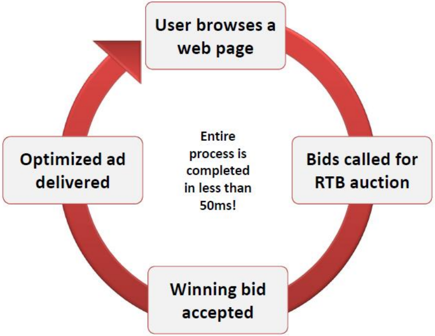
Regret

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(I_t)\right] - \min_{i=1,\dots,K} \sum_{t=1}^{T} \ell_t(i)$$

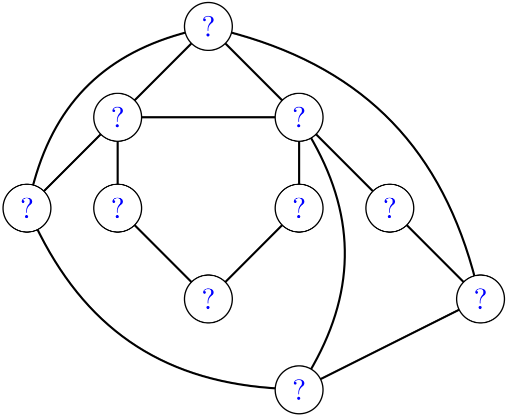
The expectation is only with respect to the player's internal randomization

Similarities between actions

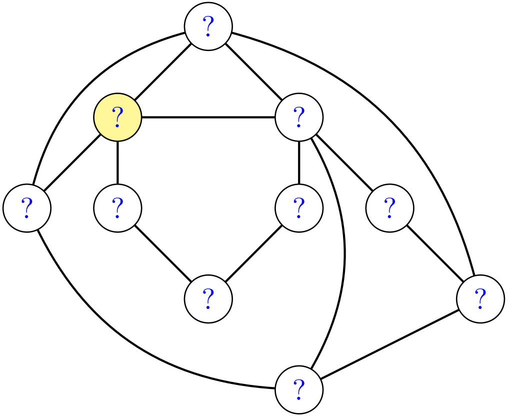




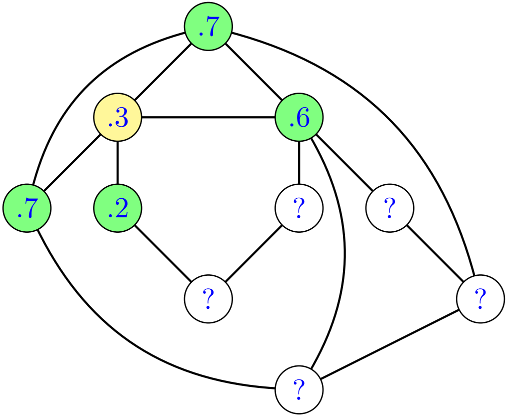
A feedback graph over actions



A feedback graph over actions



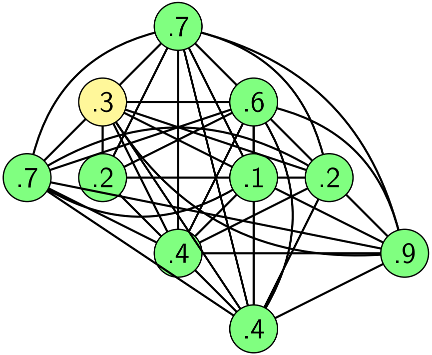
A feedback graph over actions



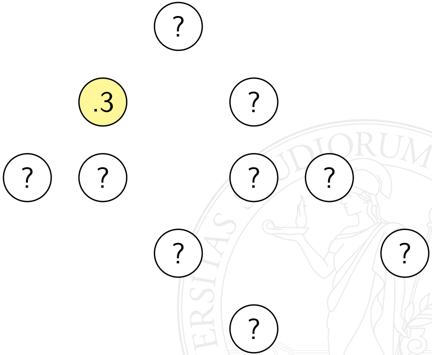
Feedback: $\ell_t(i)$ is observed iff I_t is in the neighborhood of i

Expert and bandit settings

Experts: clique



Bandits: edgeless graph



Playing on a feedback graph

Randomized player's strategy (Exp3-SET)

- $\mathbb{P}_t(I_t = i) \quad \propto \quad \exp\left(-\eta \sum_{s=1}^{t-1} \widehat{\ell}_s(i)\right) \qquad i = 1, \dots, K$

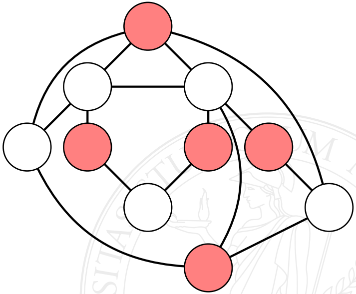
Playing on a feedback graph

Randomized player's strategy (Exp3-SET)

- $\mathbb{P}_t(I_t = i) \quad \propto \quad \exp\left(-\eta \sum_{s=1}^{t-1} \widehat{\ell}_s(i)\right) \qquad i = 1, \dots, K$
- Importance-weighted loss estimate $\widehat{\ell}_t(i) = \begin{cases} \dfrac{\ell_t(i)}{\mathbb{P}_t(\ell_t(i) \text{ observed})} & \text{if } \ell_t(i) \text{ is observed} \\ 0 & \text{otherwise} \end{cases}$

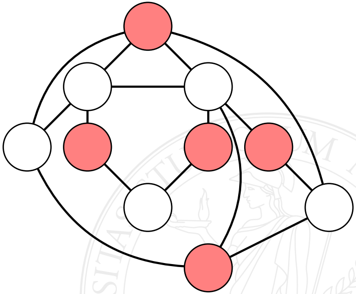
Regret bound for any feedback graph

- Let α_F be the independence number of the feedback graph



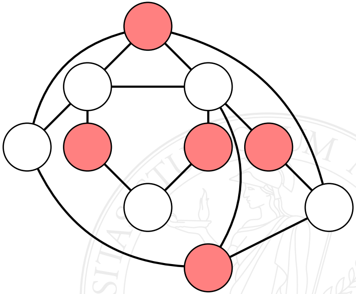
Regret bound for any feedback graph

- Let α_F be the independence number of the feedback graph
- $R_T \le \frac{\ln K}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\mathbb{E}\left[\sum_{i=1}^{K}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$



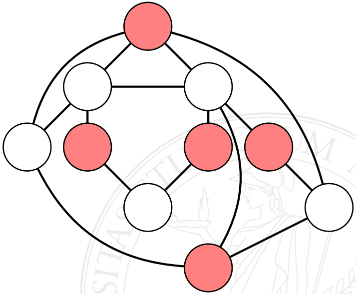
Regret bound for any feedback graph

- Let α_F be the independence number of the feedback graph
- $R_T \le \dfrac{\ln K}{\eta} + \dfrac{\eta}{2}\displaystyle\sum_{t=1}^{T}\mathbb{E}\left[\sum_{i=1}^{K}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$
- Bandit magic: $\mathbb{E}\left[\displaystyle\sum_{i=1}^{K}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right] \le \alpha_F$



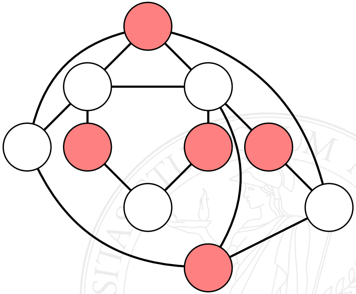
Regret bound for any feedback graph

- Let α_F be the independence number of the feedback graph
- $R_T \le \dfrac{\ln K}{\eta} + \dfrac{\eta}{2} \sum_{t=1}^{T} \mathbb{E}\left[\sum_{i=1}^{K} \mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$
- Bandit magic: $\mathbb{E}\left[\sum_{i=1}^{K} \mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right] \le \alpha_F$
- Tuning η: $R_T \stackrel{\widetilde{\mathcal{O}}}{=} \sqrt{\alpha_F T}$



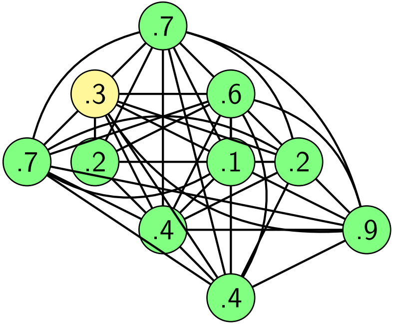
Regret bound for any feedback graph

- Let α_F be the independence number of the feedback graph
- $R_T \le \dfrac{\ln K}{\eta} + \dfrac{\eta}{2} \sum_{t=1}^{T} \mathbb{E}\left[\sum_{i=1}^{K} \mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$
- Bandit magic: $\mathbb{E}\left[\sum_{i=1}^{K} \mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right] \le \alpha_F$
- Tuning η: $R_T \overset{\widetilde{\mathcal{O}}}{=} \sqrt{\alpha_F T}$
- This bound is tight for all graphs (up to log factors)



Experts and bandits

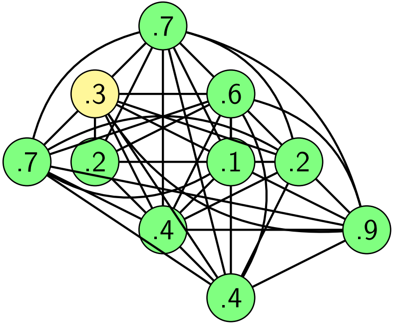
Experts: $\alpha_F = 1$



$R_T \stackrel{\widetilde{\mathcal{O}}}{=} \sqrt{T}$

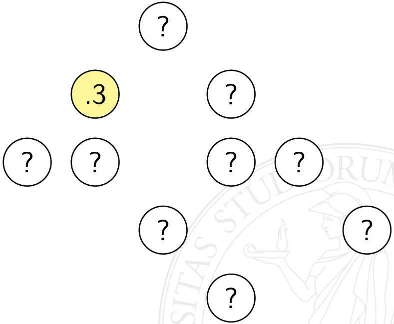
Experts and bandits

Experts: $\alpha_F = 1$



$R_T \overset{\widetilde{\mathcal{O}}}{=} \sqrt{T}$

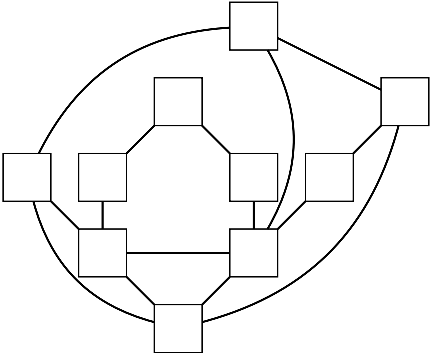
Bandits: $\alpha_F = K$



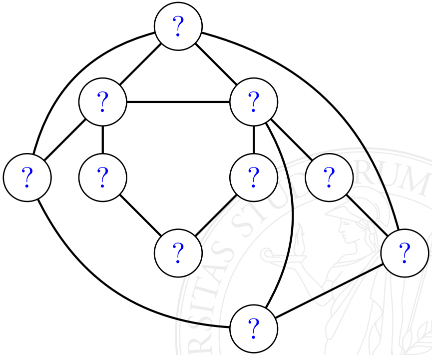
$R_T \overset{\widetilde{\mathcal{O}}}{=} \sqrt{KT}$

A protocol for multi-agent online learning

Communication graph (agents)

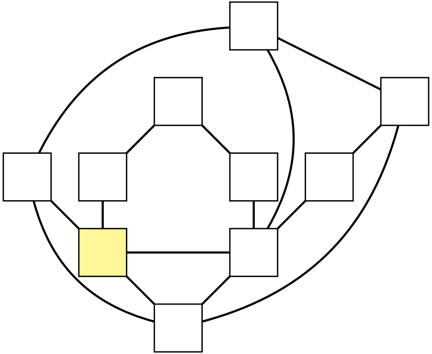


Feedback graph (actions)

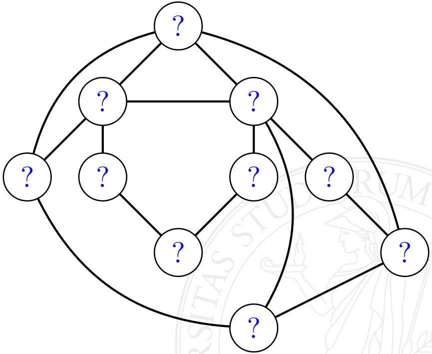


A protocol for multi-agent online learning

Communication graph (agents)



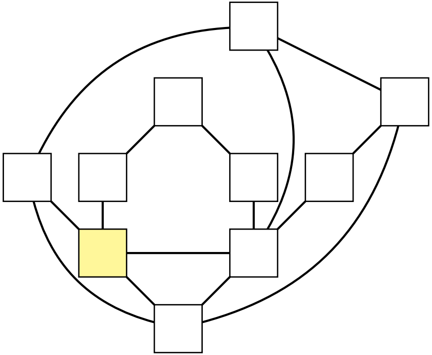
Feedback graph (actions)



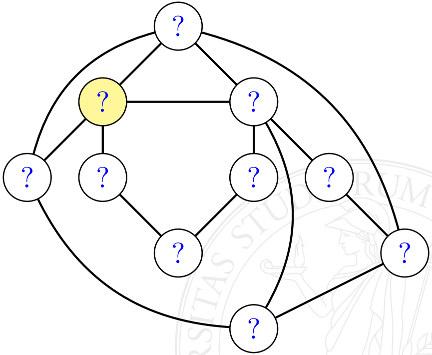
An agent v_t becomes active

A protocol for multi-agent online learning

Communication graph (agents)



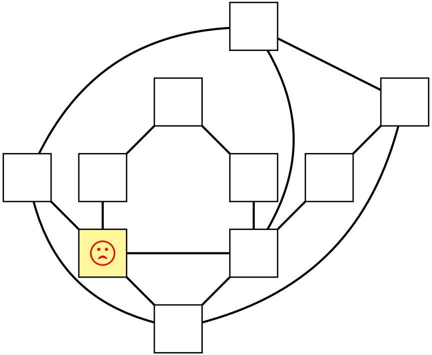
Feedback graph (actions)



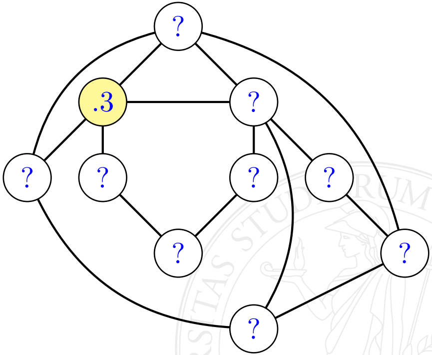
Agent v_t plays action $I_t(v_t)$

A protocol for multi-agent online learning

Communication graph (agents)



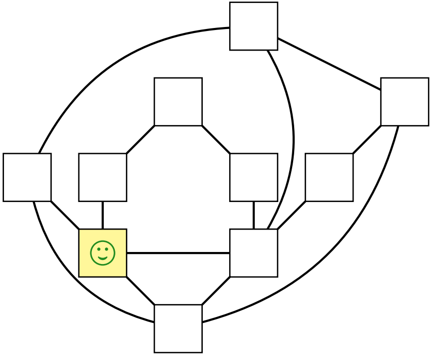
Feedback graph (actions)



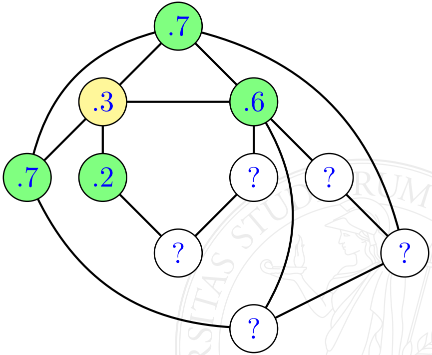
. . . and incurs loss $\ell_t(I_t(v_t))$

A protocol for multi-agent online learning

Communication graph (agents)



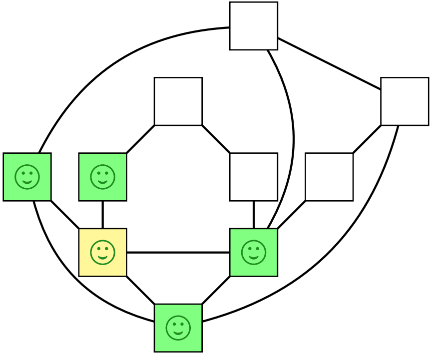
Feedback graph (actions)



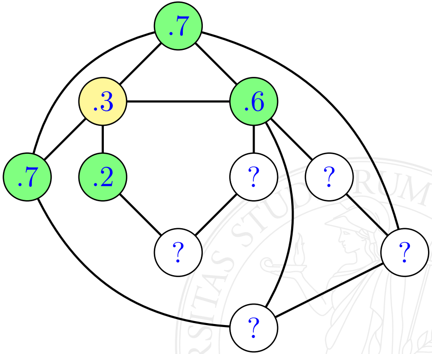
v_t observes feedback: $\ell_t(i)$ for every i in the neighborhood of I_t

A protocol for multi-agent online learning

Communication graph (agents)



Feedback graph (actions)



feedback is shared with neighbors of v_t

Network regret

- Best global action k_T^* for $\ell_1, \ell_2 \dots$ after T steps

$$k_T^* = \operatorname*{argmin}_{k=1,\dots,K} \sum_{t=1}^{T} \ell_t(k)$$

Network regret

- Best global action k_T^* for $\ell_1, \ell_2 \dots$ after T steps

$$k_T^* = \underset{k=1,\dots,K}{\operatorname{argmin}} \sum_{t=1}^{T} \ell_t(k)$$

- Assumption: each active agent v_t is drawn i.i.d. according to some fixed distribution

Network regret

- Best global action k_T^* for $\ell_1, \ell_2 \dots$ after T steps

$$k_T^* = \operatorname*{argmin}_{k=1,\dots,K} \sum_{t=1}^{T} \ell_t(k)$$

- Assumption: each active agent v_t is drawn i.i.d. according to some fixed distribution
- Network regret

$$R_T^{\text{net}} = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t\big(I_t(v_t)\big)\right] - \sum_{t=1}^{T} \ell_t(k_T^*)$$

Expectation is with respect to both agent activation and internal randomization

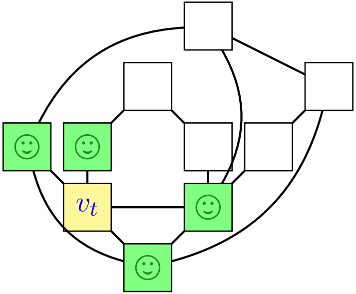
Multi-agent Exp3-SET

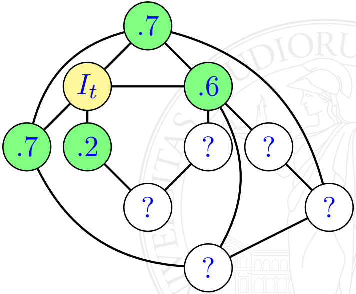
Each agent v runs Exp3-SET using

$$\widehat{\ell}_t(i,v) = \begin{cases} \dfrac{\ell_t(i)}{\mathbb{P}_t(\ell_t(i) \text{ observed by } v)} & \text{if } \ell_t(i) \text{ is observed by } v \\ 0 & \text{otherwise} \end{cases}$$

Multi-agent Exp3-SET

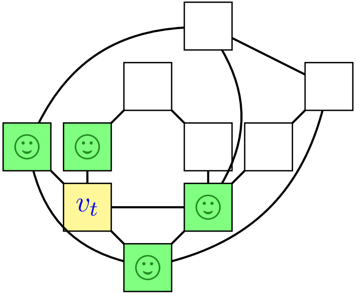
- At time t, agent v_t plays action I_t

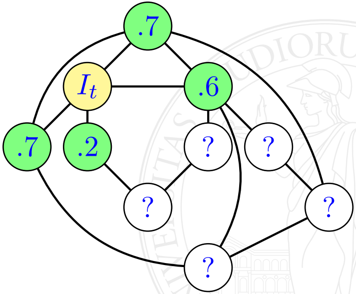




Multi-agent Exp3-SET

- At time t, agent v_t plays action I_t
- $\ell_t(i)$ may be observed by v because:
  - v is a neighbor of v_t in the communication graph (nodes are their own neighbors)
  - and i is a neighbor of $I_t(v_t)$ in the feedback graph





Old and new results (log factors ignored)

	Single-agent	Multi-agent
Experts ($F =$ clique)	$\sqrt{T}$	$\sqrt{\alpha_C T}$
Bandits ($F =$ edgeless)	$\sqrt{KT}$	$\sqrt{\alpha_C KT}$
Feedback graph	$\sqrt{\alpha_F T}$	?

Old and new results (log factors ignored)

	Single-agent	Multi-agent
Experts ($F =$ clique)	$\sqrt{T}$	$\sqrt{\alpha_C T}$
Bandits ($F =$ edgeless)	$\sqrt{KT}$	$\sqrt{\alpha_C KT}$
Feedback graph	$\sqrt{\alpha_F T}$	$\sqrt{\alpha_F \alpha_C T}$

Old and new results (log factors ignored)

	Single-agent	Multi-agent
Experts ($F =$ clique)	$\sqrt{T}$	$\sqrt{\alpha_C T}$
Bandits ($F =$ edgeless)	$\sqrt{KT}$	$\sqrt{\alpha_C KT}$
Feedback graph	$\sqrt{\alpha_F T}$	$\sqrt{\alpha^* T}$

Old and new results (log factors ignored)

	Single-agent	Multi-agent
Experts ($F =$ clique)	$\sqrt{T}$	$\sqrt{\alpha_C T}$
Bandits ($F =$ edgeless)	$\sqrt{KT}$	$\sqrt{\alpha_C KT}$
Feedback graph	$\sqrt{\alpha_F T}$	$\sqrt{\alpha^* T}$

- $\alpha^* \geq \alpha_F \alpha_C$

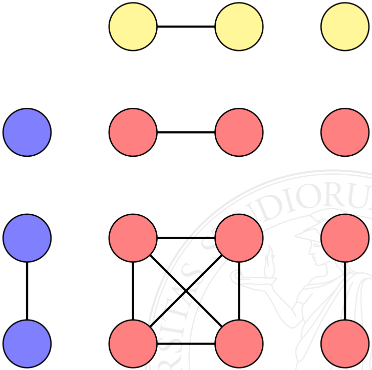
Old and new results (log factors ignored)

	Single-agent	Multi-agent
Experts ($F =$ clique)	$\sqrt{T}$	$\sqrt{\alpha_C T}$
Bandits ($F =$ edgeless)	$\sqrt{KT}$	$\sqrt{\alpha_C KT}$
Feedback graph	$\sqrt{\alpha_F T}$	$\sqrt{\alpha^* T}$

- $\alpha^* \geq \alpha_F \alpha_C$
- In general, $\alpha^* \approx \alpha_F \alpha_C$ ($\alpha^* \gg \alpha_F \alpha_C$ only in pathological cases)

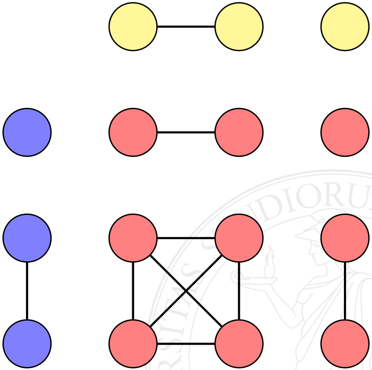
The strong product between graphs

α^* is the independence number of the strong product between the communication graph G_C and the feedback graph G_F



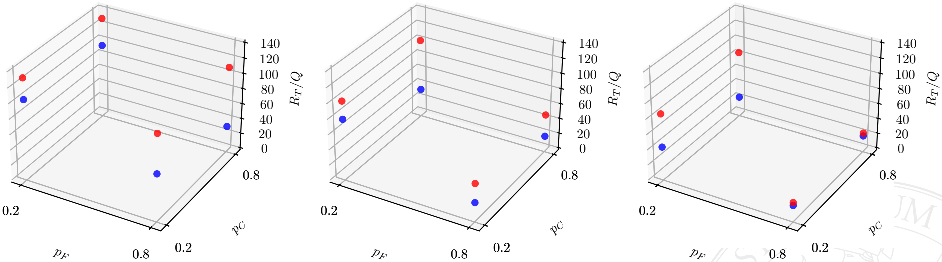
The strong product between graphs

α^* is the independence number of the strong product between the communication graph G_C and the feedback graph G_F



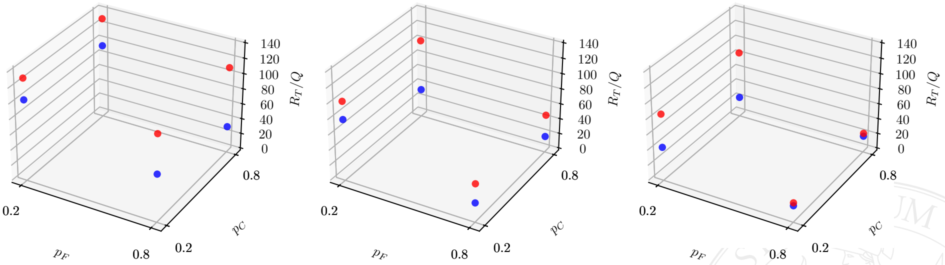
Regret bound is tight (up to log factors) for most pairs F, C

Experiments against a baseline ignoring the communication graph



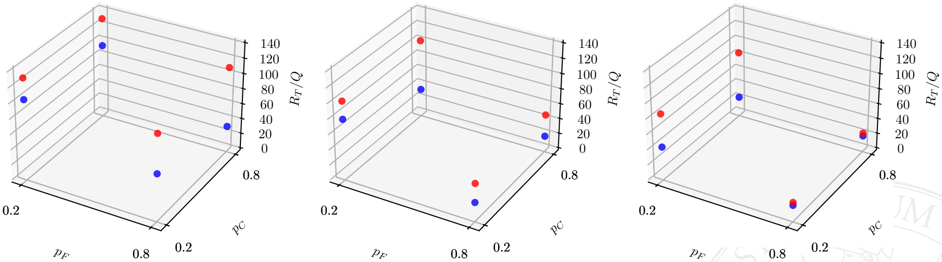
- Exp3-SET (blue dots) is never worse than the baseline (red dots)

Experiments against a baseline ignoring the communication graph



- Exp3-SET (blue dots) is never worse than the baseline (red dots)
- The performance of the baseline remains constant when p_C varies in $\{0.2, 0.8\}$. On the other hand, Exp3-SET is worse when G_C is sparse because α^* increases

Experiments against a baseline ignoring the communication graph



- Exp3-SET (blue dots) is never worse than the baseline (red dots)
- The performance of the baseline remains constant when p_C varies in $\{0.2, 0.8\}$. On the other hand, Exp3-SET is worse when G_C is sparse because α^* increases
- The performance of both algorithms is worse when G_F is sparse because α^* increases

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \ldots$

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \dots$

1. Active agent v_t is drawn i.i.d. according to some fixed distribution

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \dots$

1. Active agent v_t is drawn i.i.d. according to some fixed distribution
2. v_t predicts using the current model $\boldsymbol{x}_t(v_t) \in \mathcal{X}$ and incurs a convex loss $\ell_t(\boldsymbol{x}_t(v_t))$

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \ldots$

1. Active agent v_t is drawn i.i.d. according to some fixed distribution
2. v_t predicts using the current model $\boldsymbol{x}_t(v_t) \in \mathcal{X}$ and incurs a convex loss $\ell_t(\boldsymbol{x}_t(v_t))$
3. v_t receives gradient $\boldsymbol{g}_t = \nabla \ell_t(\boldsymbol{x}_t(v_t))$

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \ldots$

1. Active agent v_t is drawn i.i.d. according to some fixed distribution
2. v_t predicts using the current model $\boldsymbol{x}_t(v_t) \in \mathcal{X}$ and incurs a convex loss $\ell_t(\boldsymbol{x}_t(v_t))$
3. v_t receives gradient $\boldsymbol{g}_t = \nabla \ell_t(\boldsymbol{x}_t(v_t))$
4. ...and sends $\boldsymbol{A}_{v_t,v}^{-1} \boldsymbol{g}_t$ to every other agent v

Multi-agent online convex optimization

Model space: convex and closed $\mathcal{X} \subset \mathbb{R}^d$

For $t = 1, 2, \ldots$

1. Active agent v_t is drawn i.i.d. according to some fixed distribution
2. v_t predicts using the current model $\boldsymbol{x}_t(v_t) \in \mathcal{X}$ and incurs a convex loss $\ell_t(\boldsymbol{x}_t(v_t))$
3. v_t receives gradient $\boldsymbol{g}_t = \nabla \ell_t(\boldsymbol{x}_t(v_t))$
4. ... and sends $\boldsymbol{A}^{-1}_{v_t,v}\,\boldsymbol{g}_t$ to every other agent v

$\boldsymbol{A}$ is a $N \times N$ positive definite matrix of task interaction coefficients

Multi-task multi-agent online convex optimization

Multi-task multi-agent online convex optimization

- Network regret for single-task online convex optimization:

$$R_T^{\text{net}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \min_{\boldsymbol{u}\in\mathcal{X}} \sum_{t=1}^{T} \ell_t(\boldsymbol{u})$$

Multi-task multi-agent online convex optimization

- Network regret for single-task online convex optimization:

$$R_T^{\text{net}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \min_{\boldsymbol{u} \in \mathcal{X}} \sum_{t=1}^{T} \ell_t(\boldsymbol{u})$$

- All agents are learning on the same loss sequence, irrespective of their activation sequence

Multi-task multi-agent online convex optimization

- Network regret for single-task online convex optimization:

$$R_T^{\text{net}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \min_{\boldsymbol{u}\in\mathcal{X}} \sum_{t=1}^{T} \ell_t(\boldsymbol{u})$$

- All agents are learning on the same loss sequence, irrespective of their activation sequence
- Multi-task regret:

$$R_T^{\text{mt}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \sum_{v=1}^{N} \underbrace{\min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t\,:\,v_t=v} \ell_t(\boldsymbol{x})}_{\text{best local model}}$$

Multi-task multi-agent online convex optimization

- Network regret for single-task online convex optimization:

$$R_T^{\mathrm{net}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \min_{\boldsymbol{u}\in\mathcal{X}} \sum_{t=1}^{T} \ell_t(\boldsymbol{u})$$

- All agents are learning on the same loss sequence, irrespective of their activation sequence
- Multi-task regret:

$$R_T^{\mathrm{mt}} = \sum_{t=1}^{T} \ell_t(\boldsymbol{x}_t(v_t)) - \sum_{v=1}^{N} \underbrace{\min_{\boldsymbol{x}\in\mathcal{X}} \sum_{t\,:\,v_t=v} \ell_t(\boldsymbol{x})}_{\text{best local model}}$$

- Each agent learns on its local loss sequence, defined by their activation sequence

Single-agent online gradient descent

Follow the Regularized Leader (FTRL)

Single-agent online gradient descent

Follow the Regularized Leader (FTRL)

- $\boldsymbol{x}_t = \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} \left(\frac{1}{2} \|\boldsymbol{x}\|^2 + \eta \, \boldsymbol{x}^\top \sum_{s=1}^{t-1} \boldsymbol{g}_s \right)$

Single-agent online gradient descent

Follow the Regularized Leader (FTRL)

- $\boldsymbol{x}_t = \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{argmin}} \left(\frac{1}{2} \|\boldsymbol{x}\|^2 + \eta\, \boldsymbol{x}^\top \sum_{s=1}^{t-1} \boldsymbol{g}_s \right)$
- Regret bound: $R_T \le \frac{D^2}{2\eta} + \frac{\eta}{2} L_2^2 T = DL_2\sqrt{T}$

Single-agent online gradient descent

Follow the Regularized Leader (FTRL)

- $\boldsymbol{x}_t = \underset{\boldsymbol{x} \in \mathcal{X}}{\operatorname{argmin}} \left(\frac{1}{2} \|\boldsymbol{x}\|^2 + \eta \, \boldsymbol{x}^\top \sum_{s=1}^{t-1} \boldsymbol{g}_s \right)$
- Regret bound: $R_T \le \frac{D^2}{2\eta} + \frac{\eta}{2} L_2^2 T = D L_2 \sqrt{T}$

D is the diameter of $\mathcal{X}$

L_2 is a uniform bound on the Lipschitz constant of ℓ_t

Multi-task multi-agent FTRL

- Fix any positive definite matrix A of task interaction coefficients

Multi-task multi-agent FTRL

- Fix any positive definite matrix $\boldsymbol{A}$ of task interaction coefficients
- $\boldsymbol{X}_t = \begin{bmatrix} \vdots \\ \boldsymbol{x}_t(v)^\top \\ \vdots \end{bmatrix}$ is $N \times d$ matrix of local models

Multi-task multi-agent FTRL

- Fix any positive definite matrix $\boldsymbol{A}$ of task interaction coefficients
- $\boldsymbol{X}_t = \begin{bmatrix} \vdots \\ \boldsymbol{x}_t(v)^\top \\ \vdots \end{bmatrix}$ is $N \times d$ matrix of local models
- Matrix FTRL: $\boldsymbol{X}_t = \operatorname*{argmin}_{\boldsymbol{X}} \left(\frac{1}{2} \|\boldsymbol{X}\|_{\boldsymbol{A}}^2 + \eta \sum_{s=1}^{t-1} \langle \boldsymbol{X}, \boldsymbol{G}_s \rangle \right)$

Multi-task multi-agent FTRL

- Fix any positive definite matrix $\boldsymbol{A}$ of task interaction coefficients
- $\boldsymbol{X}_t = \begin{bmatrix} \vdots \\ \boldsymbol{x}_t(v)^\top \\ \vdots \end{bmatrix}$ is $N \times d$ matrix of local models
- Matrix FTRL: $\boldsymbol{X}_t = \operatorname*{argmin}_{\boldsymbol{X}} \left(\frac{1}{2} \|\boldsymbol{X}\|_{\boldsymbol{A}}^2 + \eta \sum_{s=1}^{t-1} \langle \boldsymbol{X}, \boldsymbol{G}_s \rangle \right)$
- $\boldsymbol{G}_s$ is $N \times d$ matrix with only one non-zero row, $\boldsymbol{g}_s = \nabla \ell_s(\boldsymbol{x}_s(v_s))$

Multi-task multi-agent FTRL

- Fix any positive definite matrix $\boldsymbol{A}$ of task interaction coefficients
- $\boldsymbol{X}_t = \begin{bmatrix} \vdots \\ \boldsymbol{x}_t(v)^\top \\ \vdots \end{bmatrix}$ is $N \times d$ matrix of local models
- Matrix FTRL: $\boldsymbol{X}_t = \operatorname*{argmin}_{\boldsymbol{X}} \left(\frac{1}{2} \|\boldsymbol{X}\|_{\boldsymbol{A}}^2 + \eta \sum_{s=1}^{t-1} \langle \boldsymbol{X}, \boldsymbol{G}_s \rangle \right)$
- $\boldsymbol{G}_s$ is $N \times d$ matrix with only one non-zero row, $\boldsymbol{g}_s = \nabla \ell_s(\boldsymbol{x}_s(v_s))$
- $\|\boldsymbol{X}\|_{\boldsymbol{A}}^2 = \operatorname{tr}(\boldsymbol{A}\boldsymbol{X}\boldsymbol{X}^\top)$

Regret bound

- Matrix FTRL regret bound: $R_T^{\mathrm{mt}} \le \frac{\|\boldsymbol{U}\|_{\boldsymbol{A}}^2}{2\eta} + \frac{\eta}{2} L_2^2 \sum_{t=1} A_{v_t,v_t}^{-1}$

Regret bound

- Matrix FTRL regret bound: $R_T^{\mathrm{mt}} \le \dfrac{\|\boldsymbol{U}\|_{\boldsymbol{A}}^2}{2\eta} + \dfrac{\eta}{2} L_2^2 \sum_{t=1} A_{v_t,v_t}^{-1}$
- $\boldsymbol{U}$ is $N \times d$ matrix of best local models $\boldsymbol{x}^*(v) = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} \sum_{t\,:\,v_t=v} \ell_t(\boldsymbol{x})$

Regret bound

- Matrix FTRL regret bound: $R_T^{\mathrm{mt}} \le \frac{\|\boldsymbol{U}\|_{\boldsymbol{A}}^2}{2\eta} + \frac{\eta}{2} L_2^2 \sum_{t=1} A_{v_t,v_t}^{-1}$
- $\boldsymbol{U}$ is $N \times d$ matrix of best local models $\boldsymbol{x}^*(v) = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} \sum_{t\,:\,v_t=v} \ell_t(\boldsymbol{x})$
- $\boldsymbol{A} = \boldsymbol{I}_N$ (no interaction) implies $R_T^{\mathrm{mt}} \le D L_2 \sqrt{NT}$

Parametric choice of A

- Set $\boldsymbol{A} = (1+b)\boldsymbol{I}_N - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$ for $b > 0$

Parametric choice of A

- Set $\boldsymbol{A} = (1+b)\boldsymbol{I}_N - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$ for $b > 0$
- Assume best local models $\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)$ satisfy $\|\boldsymbol{x}^*(v)\|_2 \le D$ and

$$\mathrm{Var}\Big(\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)\Big) = \frac{1}{N-1} \sum_{v=1}^{N} \|\boldsymbol{x}^*(v) - \overline{\boldsymbol{x}}\|_2^2 \le (\sigma D)^2$$

Parametric choice of A

- Set $\boldsymbol{A} = (1+b)\boldsymbol{I}_N - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$ for $b > 0$
- Assume best local models $\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)$ satisfy $\|\boldsymbol{x}^*(v)\|_2 \le D$ and

$$\mathrm{Var}\Big(\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)\Big) = \frac{1}{N-1}\sum_{v=1}^{N} \|\boldsymbol{x}^*(v) - \overline{\boldsymbol{x}}\|_2^2 \le (\sigma D)^2$$

- $R_T^{\mathrm{mt}} \le D L_2 \sqrt{2\big(1 + (N-1)\sigma^2\big) T}$

Parametric choice of A

- Set $\boldsymbol{A} = (1+b)\boldsymbol{I}_N - \frac{b}{N}\mathbf{1}\mathbf{1}^\top$ for $b > 0$
- Assume best local models $\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)$ satisfy $\|\boldsymbol{x}^*(v)\|_2 \le D$ and

$$\mathrm{Var}\Big(\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)\Big) = \frac{1}{N-1}\sum_{v=1}^{N} \|\boldsymbol{x}^*(v) - \overline{\boldsymbol{x}}\|_2^2 \le (\sigma D)^2$$

- $R_T^{\mathrm{mt}} \le DL_2\sqrt{2(1+(N-1)\sigma^2)T}$
- Always better than the no-interaction bound $DL_2\sqrt{NT}$ for $\sigma^2 < 1$

Parametric choice of A

- Set $\boldsymbol{A} = (1+b)\boldsymbol{I}_N - \dfrac{b}{N}\boldsymbol{1}\boldsymbol{1}^\top$ for $b > 0$
- Assume best local models $\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)$ satisfy $\|\boldsymbol{x}^*(v)\|_2 \le D$ and

$$\mathrm{Var}\Big(\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)\Big) = \frac{1}{N-1}\sum_{v=1}^{N} \|\boldsymbol{x}^*(v) - \overline{\boldsymbol{x}}\|_2^2 \le (\sigma D)^2$$

- $R_T^{\mathrm{mt}} \le DL_2\sqrt{2\big(1+(N-1)\sigma^2\big)T}$
- Always better than the no-interaction bound $DL_2\sqrt{NT}$ for $\sigma^2 < 1$

Matching lower bound

Under the same conditions on $\boldsymbol{x}^*(1), \dots, \boldsymbol{x}^*(N)$, any online algorithm satisfies

$$R_T^{\mathrm{mt}} \ge \frac{1}{4} DL_2\sqrt{2\big(1+(N-1)\sigma^2\big)T}$$

Learning the variance

- If $\sigma^2 \geq 1$, then we set $b = 0$ and run independent FTRL

Learning the variance

- If $\sigma^2 \geq 1$, then we set $b = 0$ and run independent FTRL
- To learn $\sigma^2 \in [0, 1]$ we can run N instances of Matrix FTRL on a $\frac{1}{N}$-grid

Learning the variance

- If $\sigma^2 \geq 1$, then we set $b = 0$ and run independent FTRL
- To learn $\sigma^2 \in [0,1]$ we can run N instances of Matrix FTRL on a $\frac{1}{N}$-grid
- These instances are aggregated via prediction with expert advice (Hedge algorithm)

Learning the variance

- If $\sigma^2 \geq 1$, then we set $b = 0$ and run independent FTRL
- To learn $\sigma^2 \in [0, 1]$ we can run N instances of Matrix FTRL on a $\frac{1}{N}$-grid
- These instances are aggregated via prediction with expert advice (Hedge algorithm)
- Resulting bound

$$R_T^{\mathrm{mt}} \leq DL_2 \left(2 + \ln N + \sqrt{2N \min\{1, \sigma^2\} T} \right)$$

Experiments

- Both Euclidean (MT-OGD) and Entropic (MT-EG) variants have closed form updates.

Experiments

- Both Euclidean (MT-OGD) and Entropic (MT-EG) variants have closed form updates.
- Independent task (IT, $b = 0$) and single task (ST, $b = +\infty$) and EG on *Lenk* (σ^2 small) and *EMNIST* (σ^2 large)