GRAPHS IN MACHINE LEARNING

Michal Valko, SequeL, Inria Lille - Nord Europe

TA: **Pierre Perrault**

MVA 2018/2019 Partially based on material by Tomáš Kocák

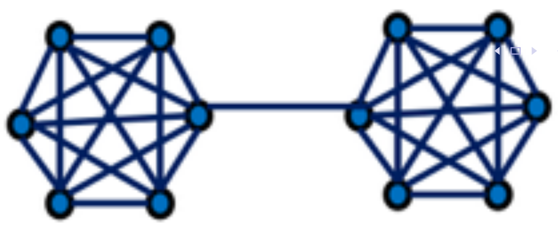
NEXT LAB SESSION

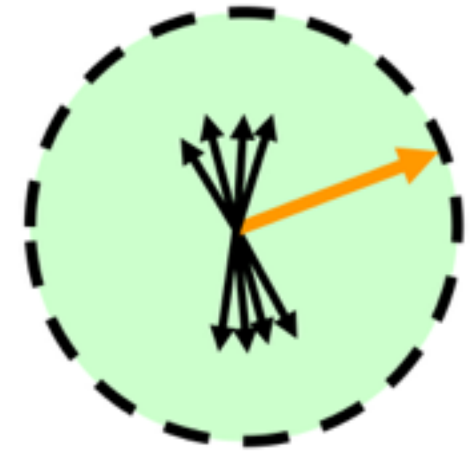
- **DL for TD2: today**
- **No class or lab (TD) next week**
- 12. 12. 2018 by Pierre Perrault
- Content: Online and scalable algorithms
  - Online face recognizer
  - Iterative label propagation
  - Online k-centers
- AR: record a video with faces
- Short written report
- Questions to piazza
- *Deadline:* 26. 12. 2018

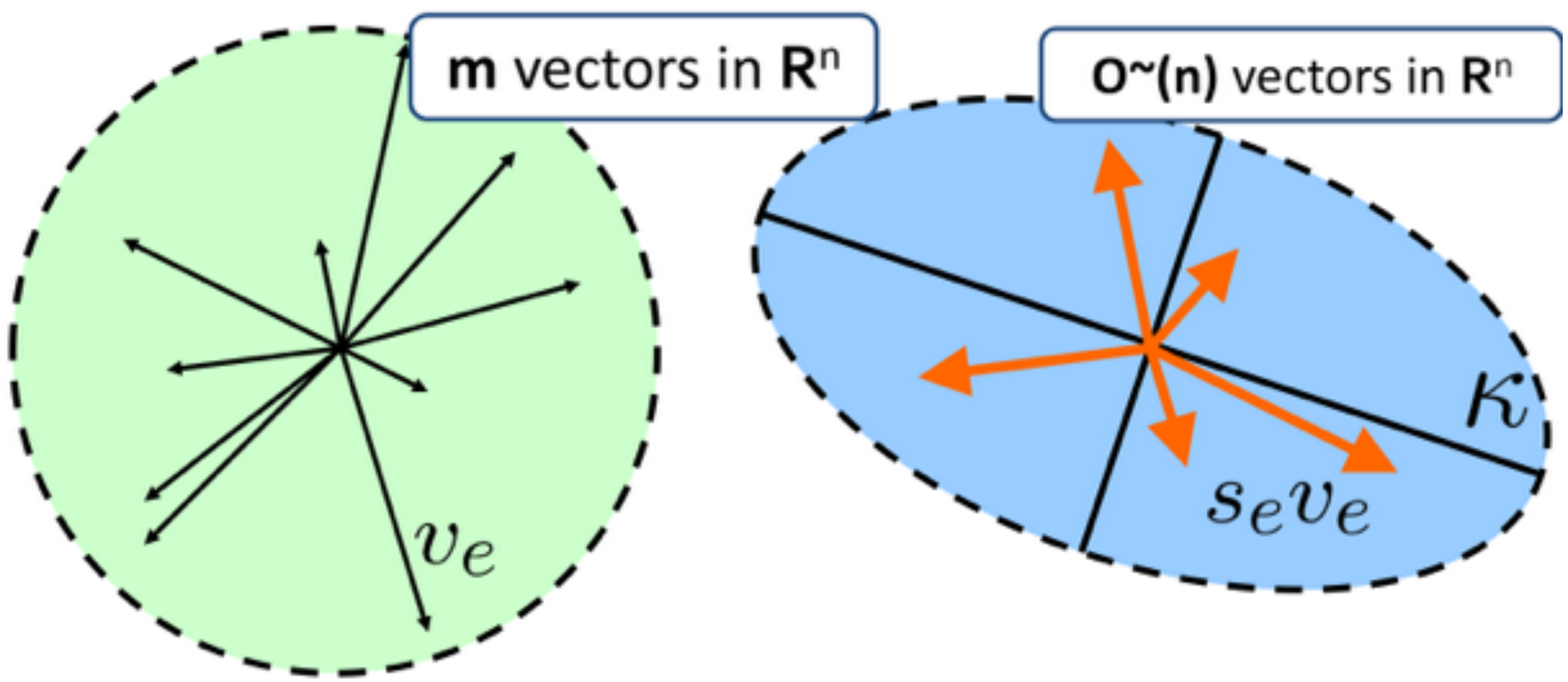
FINAL CLASS PROJECTS

- detailed description on the class website
- preferred option: you come up with the topic
- theory/implementation/review or a combination
- one or two people per project (exceptionally three)
- grade 60%: report + short presentation of the **team**
- deadlines
  - 21. 11. 2018 - strongly recommended DL for taking projects
  - TODAY 28. 11. 2018 - hard DL for taking projects
  - 07. 01. 2019 - submission of the project report
  - 11. 01. 2019 or later - project presentation
- list of suggested topics on piazza

PREVIOUS LECTURE







Questions?

MEET THE QUEEN!

What? Internships (6 months) and PhD positions (3 years)
When? From March 2019 (internships) and October 2019 (PhD)
Where? London, UK
With who? Dr. Benjamin Guedj (researcher @Inria @UCL)
What for? Invention, analysis, implementation of an agnostic learning framework through the use of the PAC-Bayesian theory
Huh? PAC-what? Check out the NIPS 2017 workshop!

`https://bguedj.github.io`

NIPS 2017 Workshop

(Almost) 50 Shades of Bayesian Learning: PAC-Bayesian trends and insights

Long Beach Convention Center, California
December 9, 2017

THIS LECTURE LAST LECTURE OF THE COURSE

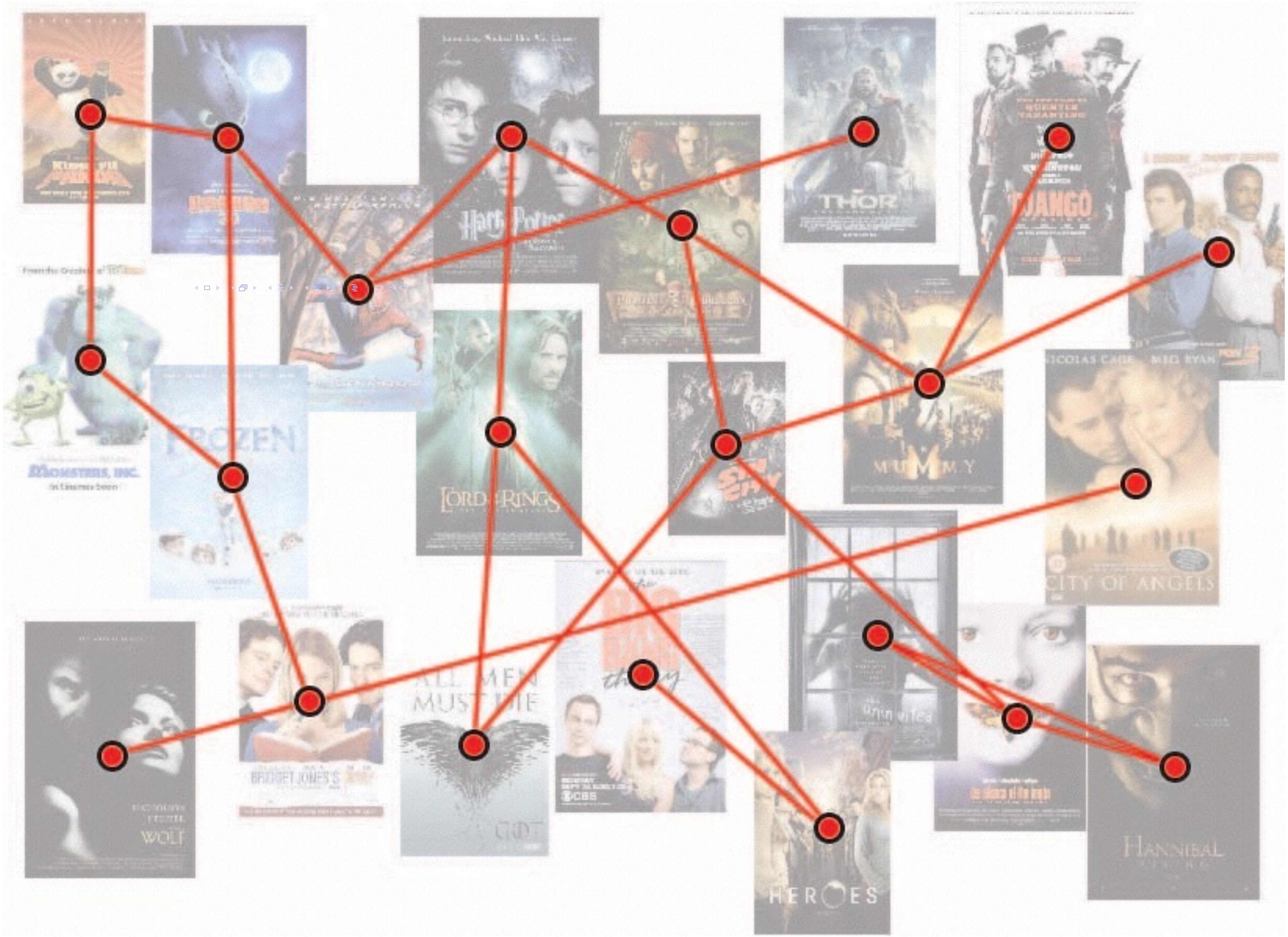
- Graph bandits
  - Spectral bandits
  - Observability graphs
  - Side information
  - Influence Maximization

RL/BANDITS ~ SEQUENTIAL DECISION-MAKING

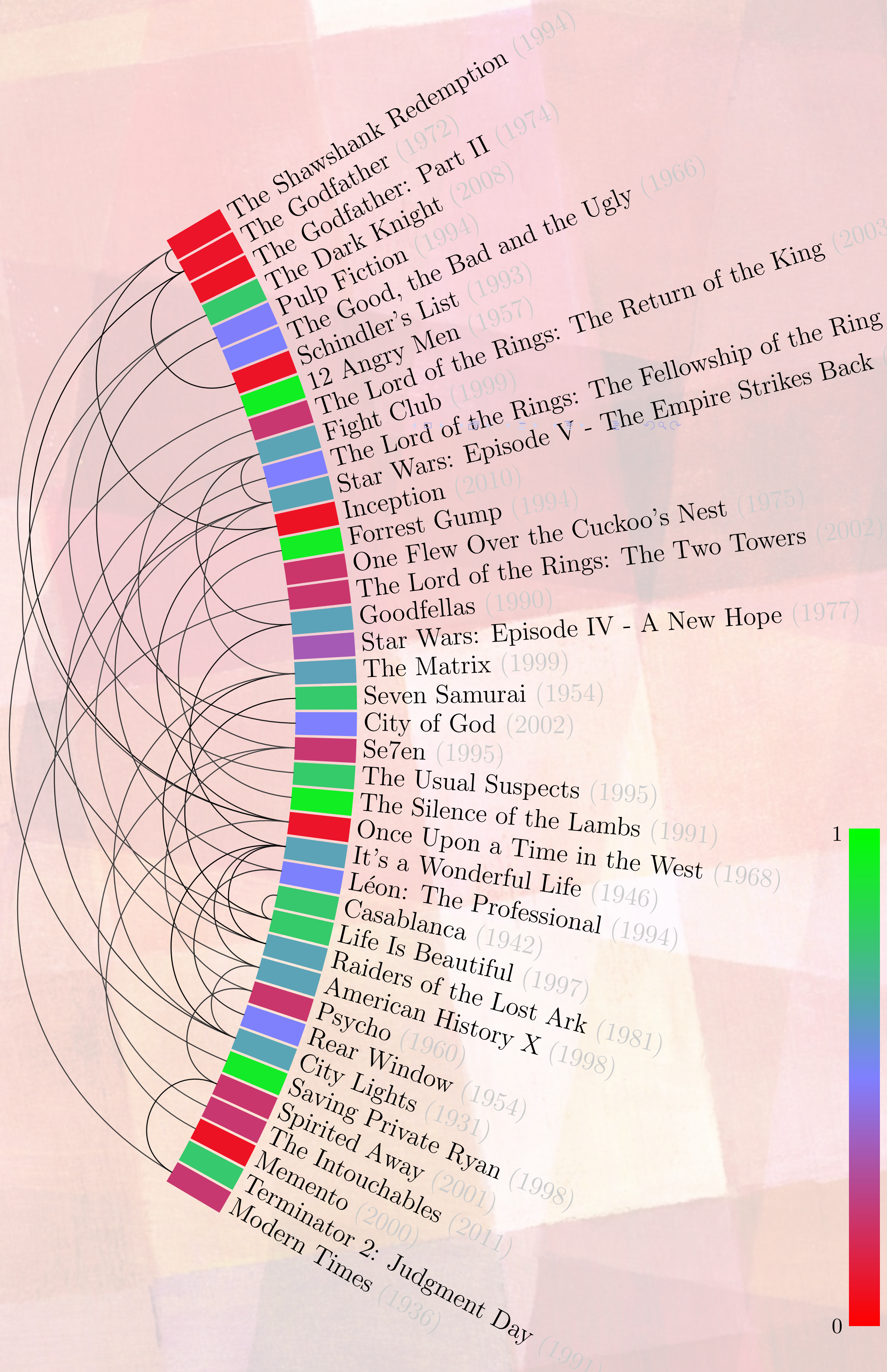
unsupervised - supervised-semisupervised-active

MULTI-ARM BANDITS IN LAS VEGAS
DECEMBER 2017

ps: several course projects are on this topic



THOR
DJANGO
MONSTERS, INC.
FROZEN
LORD RINGS
MUMMY
CITY OF ANGELS
ALL MEN MUST DIE
GOT
BRIDGET JONES'S
WOLF
HEROES
HANNIBAL



Example of a graph bandit problem

movie recommendation

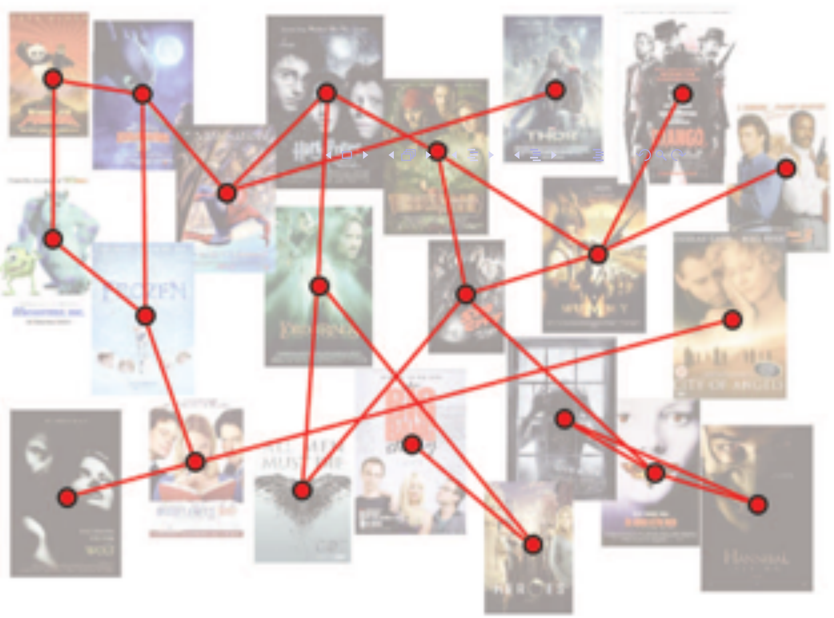
- recommend movies to a **single** user
- **goal:** maximise the sum of the ratings (minimise regret)
- good prediction after just a few steps

$$T \ll N$$

- extra information
  - ratings are **smooth** on a graph
- main question: can we learn **faster**?

GETTING REAL

Let's be lazy and ignore the structure



→

Multi-armed bandit problem!

Worst case regret (to the best fixed strategy)

$$R_T = \mathcal{O}\left(\sqrt{NT}\right)$$

#actions

#rounds

Matching lower bound (Auer, Cesa-Bianchi, Freund, Schapire 2002)

How big is N? Number of movies on **http://www.imdb.com/stats**: 4,029,967

Problem: Too many actions!

LEARNING FASTER

#actions

$$R_T = \mathcal{O}\left(\sqrt{NT}\right)$$

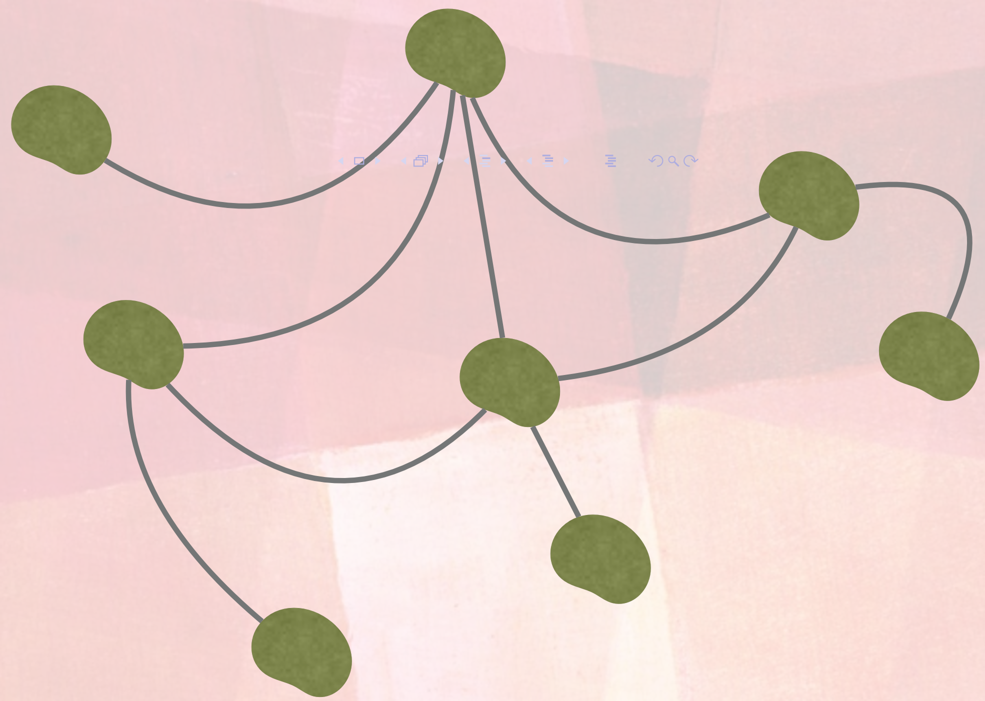
#rounds

- Arm independence is too strong and unnecessary
- Replace **N** with something much smaller
  - problem/instance/data dependent
  - example: linear bandits **N** to **D**

#dimensions

- **Today: Graph** Bandits!
  - sequential problems where **actions are nodes** on a graph
  - find strategies that replace **N** with a smaller **graph-dependent** quantity

GRAPH BANDITS: GENERAL SETUP



Every round **t** the learner

- picks a node $I_t \in [N]$
- incurs a loss ℓ_{t,I_t}
- optional feedback

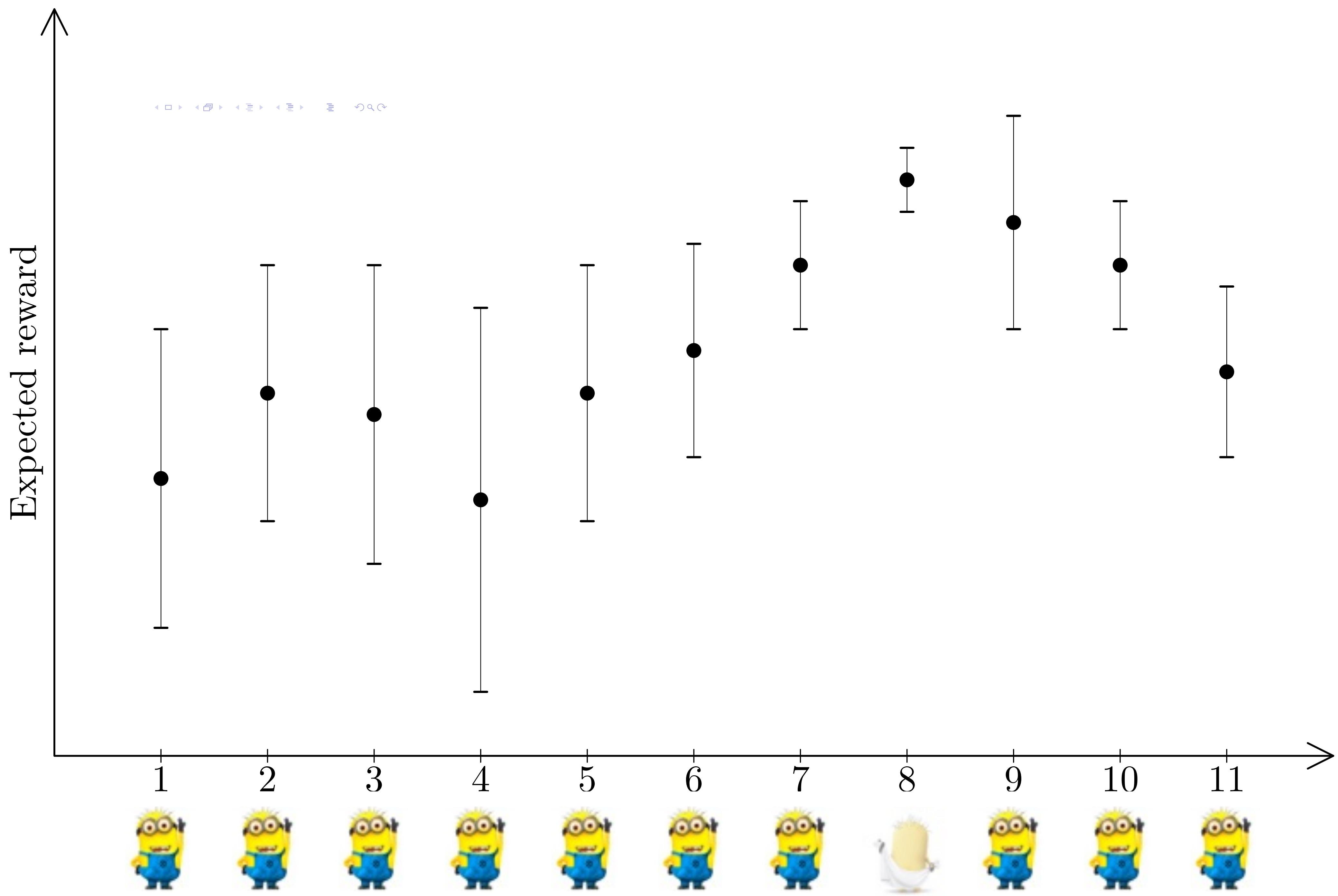
The performance is total expected regret

$$R_T = \max_{i \in [N]} \mathbb{E}\left[\sum_{t=1}^{T} (\ell_{t,I_t} - \ell_{t,i})\right]$$

Specific problems differ in

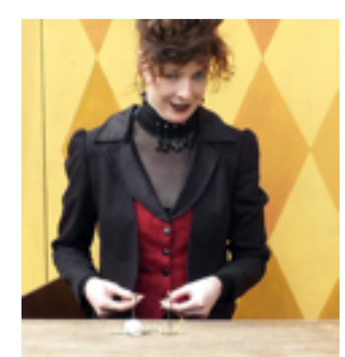
1. loss
2. feedback
3. guarantees

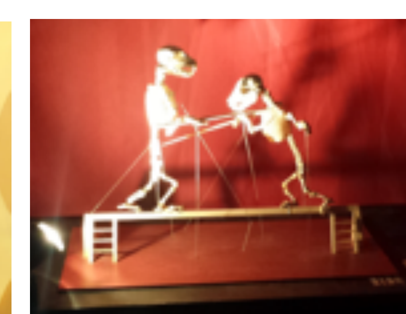
UPPER CONFIDENCE BOUND BASED ALGOS



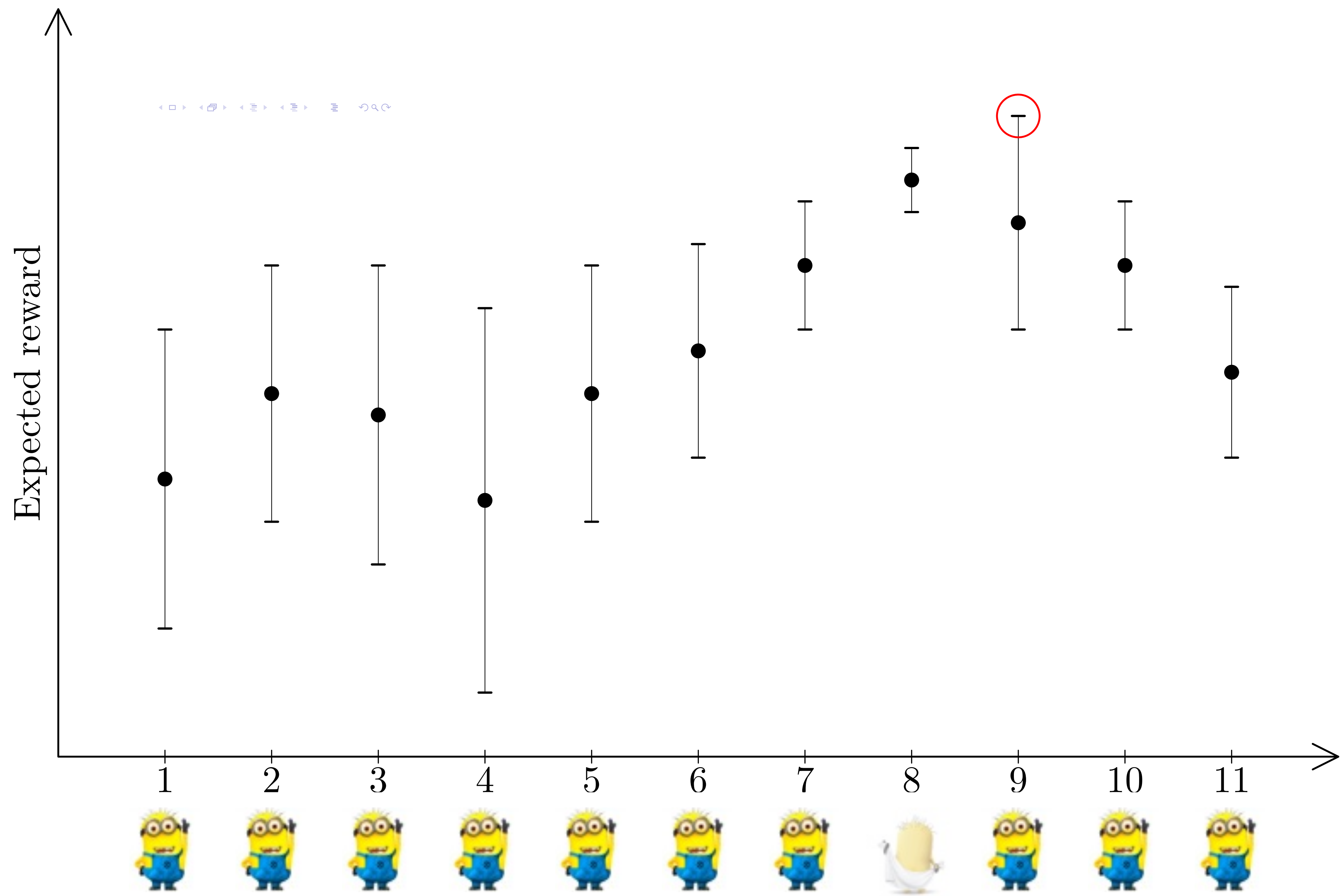
MULTI-ARM BANDITS IN CAFÉ CULTURE

Video recorded March 30th, 2015, 13h50,
Université de Lille, Susie & the Piggy Bones Band

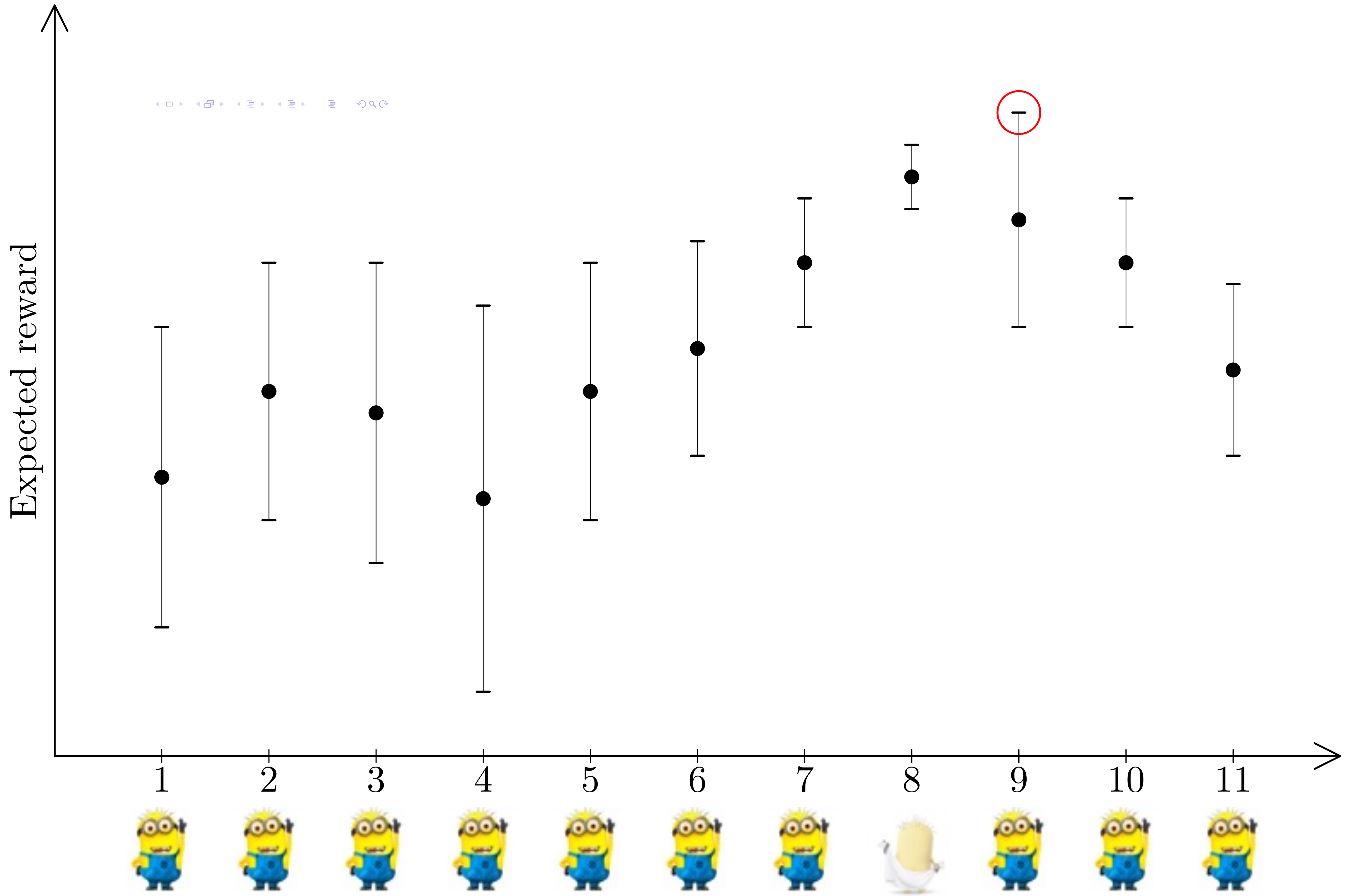




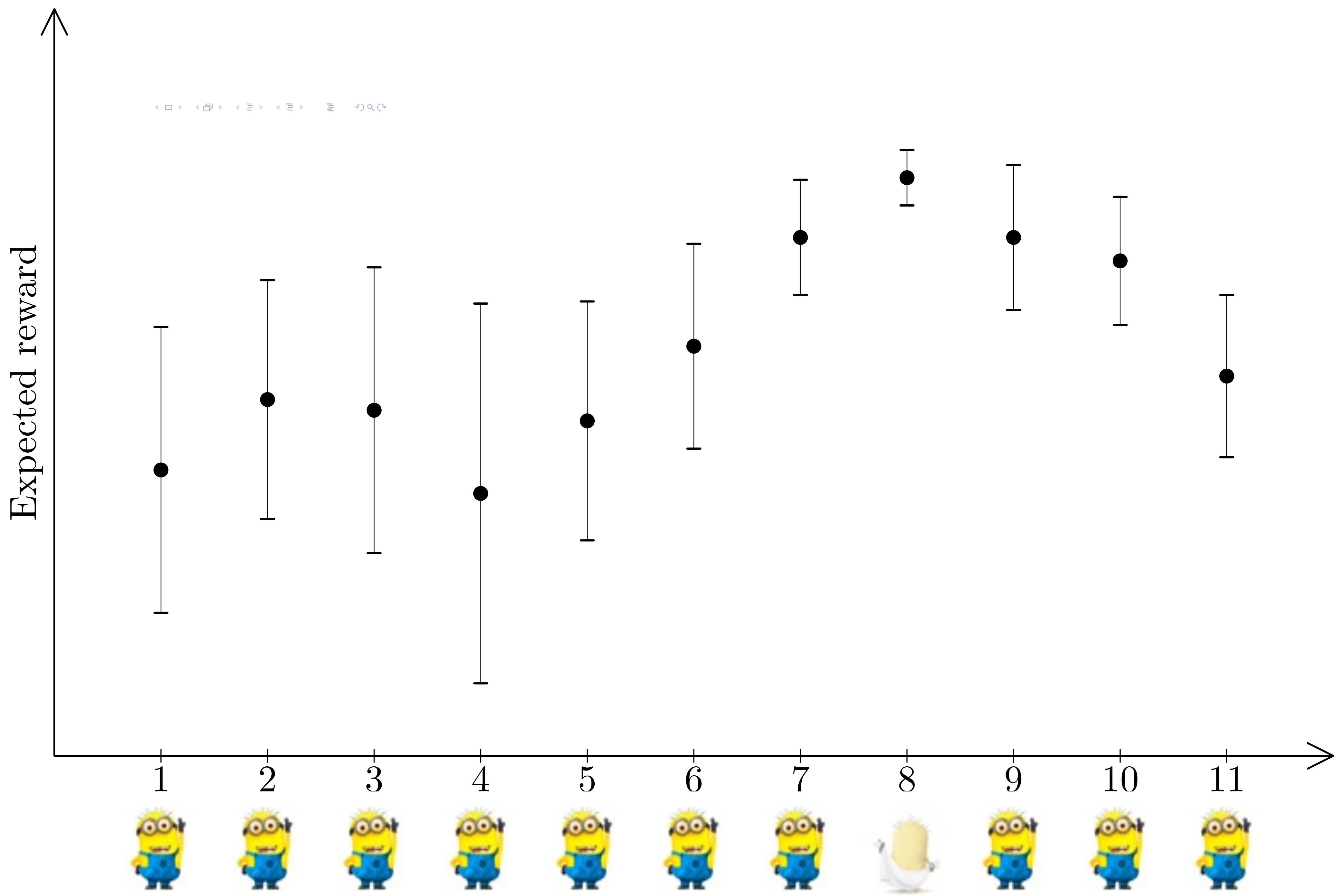
UPPER CONFIDENCE BOUND BASED ALGOS



UPPER CONFIDENCE BOUND BASED ALGOS

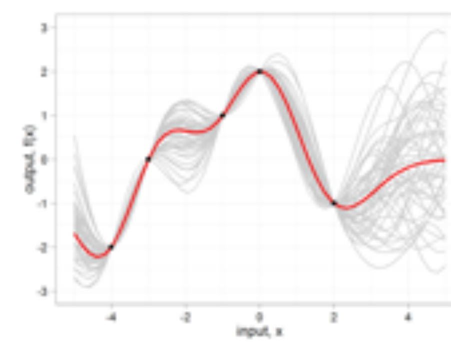


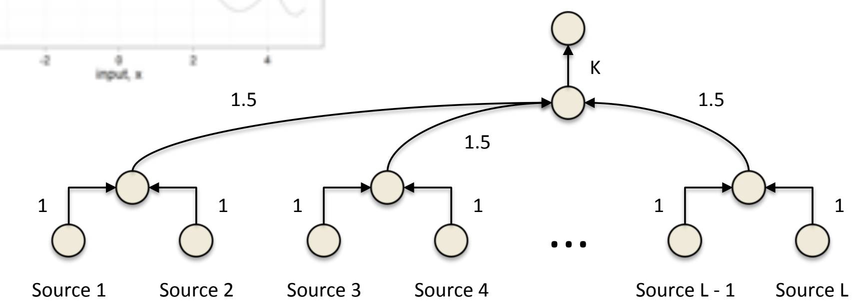
UPPER CONFIDENCE BOUND BASED ALGOS

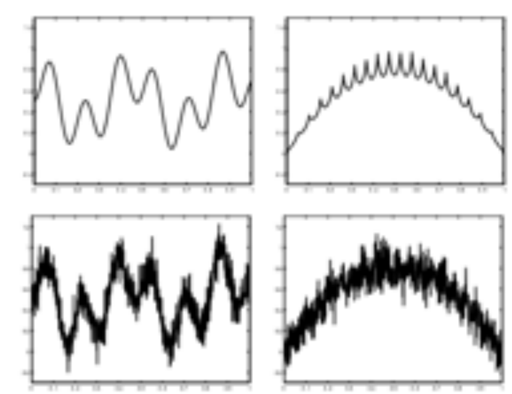


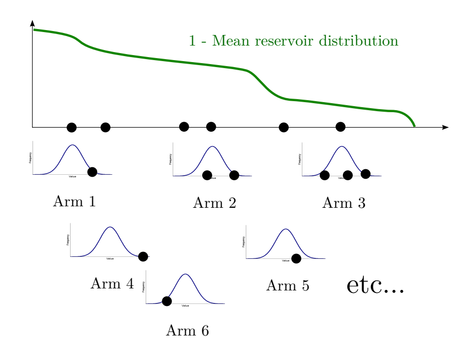
STRUCTURES IN BANDIT PROBLEMS

- GRAPHS
- KERNELS
- POLYMATROIDS
- BLACK-BOX FUNCTIONS
- STRUCTURES WITHOUT TOPOLOGY
- …









SPECIFIC GRAPH BANDIT SETTINGS

informatics mathematics Inria

Survey: http://researchers.lille.inria.fr/~valko/hp/publications/valko2016bandits.pdf

smoothness
spectral bandits
$$R_T = \widetilde{\mathcal{O}}\left(d\sqrt{T \ln T}\right)$$
#relevant eigenvectors

side observations
on graphs
$$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha T \ln N}\right)$$
independence number

influence maximisation
revealing bandits
$$R_T = \widetilde{\mathcal{O}}\left(\sqrt{r_* T D_*}\right)$$
detectable dimension

noisy side observations
on graphs
$$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha^\star T \ln N}\right)$$
effective independence number

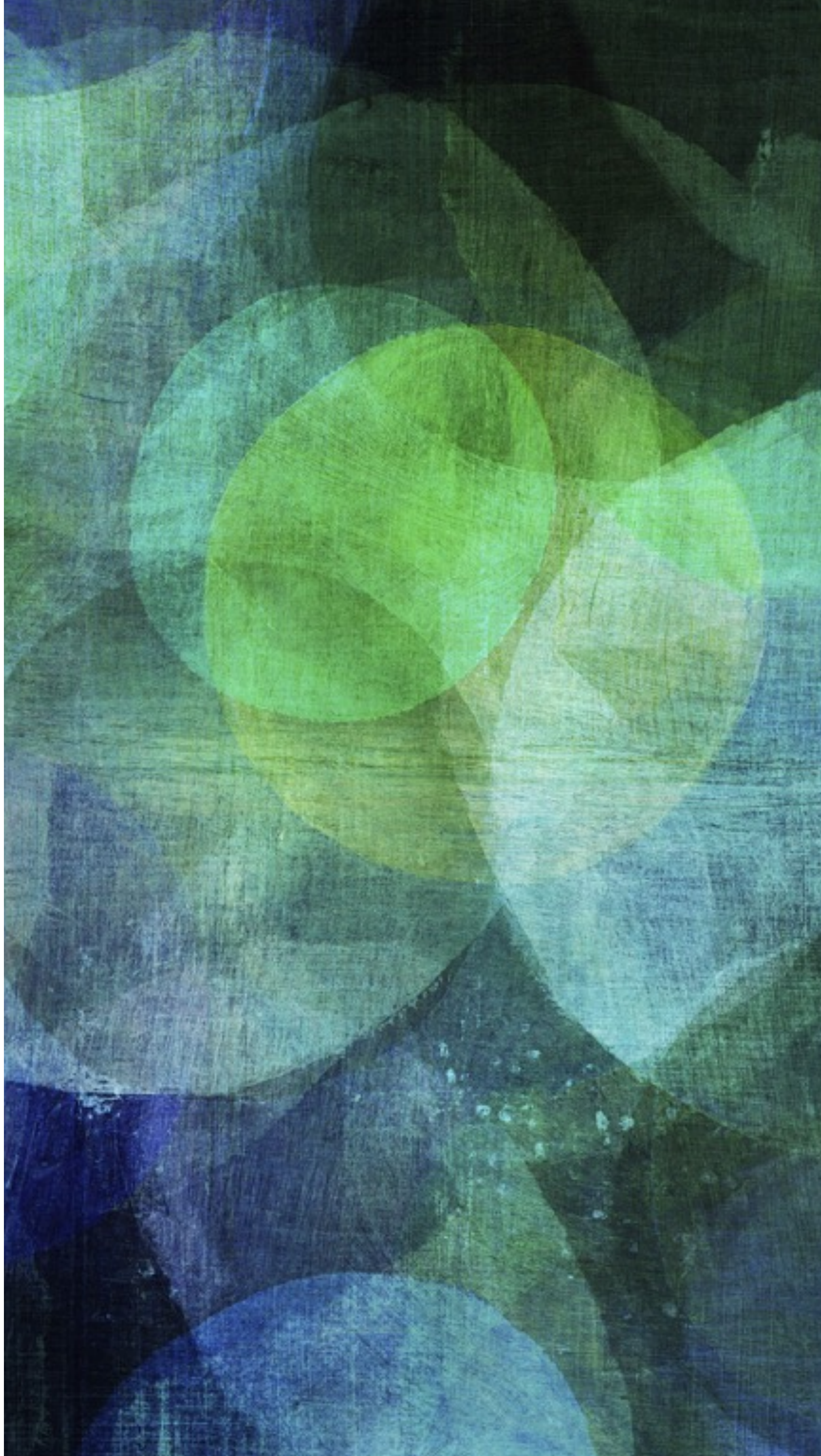
Kocák, Neu, MV, Munos: **Efficient learning by implicit exploration in bandit problems with side observations**, NIPS 2014

Kocák, Neu, MV: **Online learning with Erdos-Rényi side-observation graphs** UAI 2016

Kocák, Neu, MV: **Online learning with noisy side observations**, AISTATS 2016

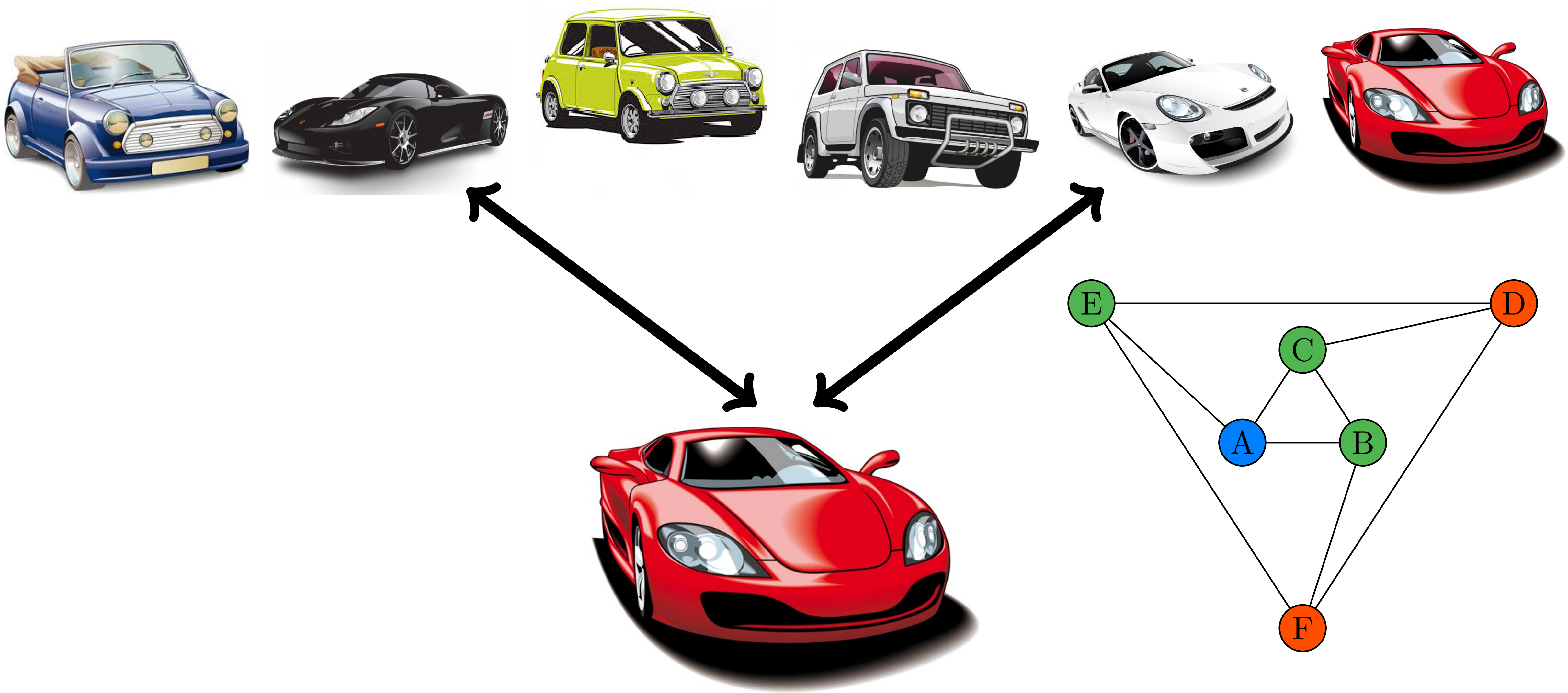
GRAPH BANDITS WITH SIDE OBSERVATIONS

exploiting free observations from neighbouring nodes



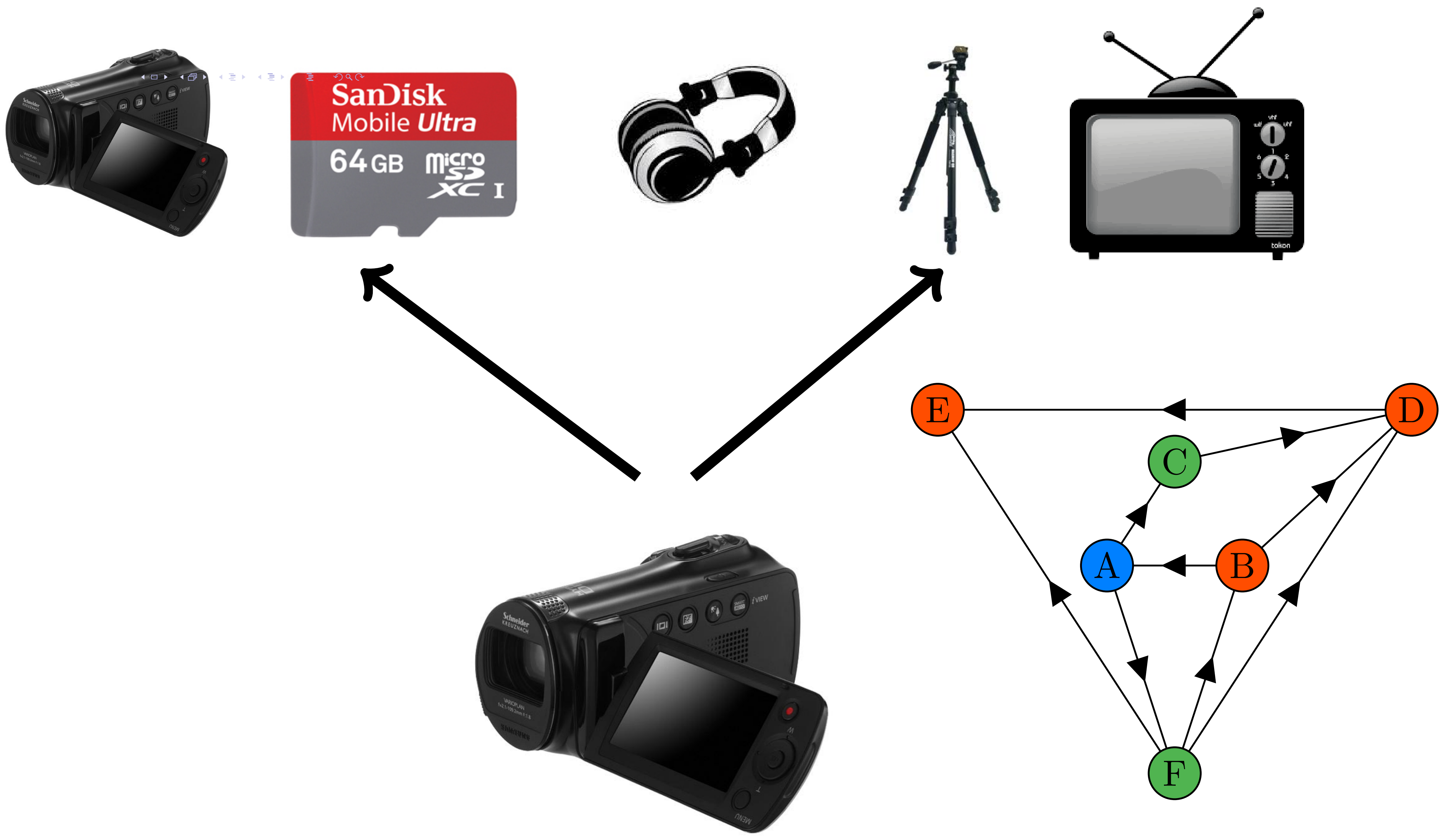
SIDE OBSERVATIONS: UNDIRECTED

Example 1: undirected observations



SIDE OBSERVATIONS: DIRECTED

Example 2: Directed observation

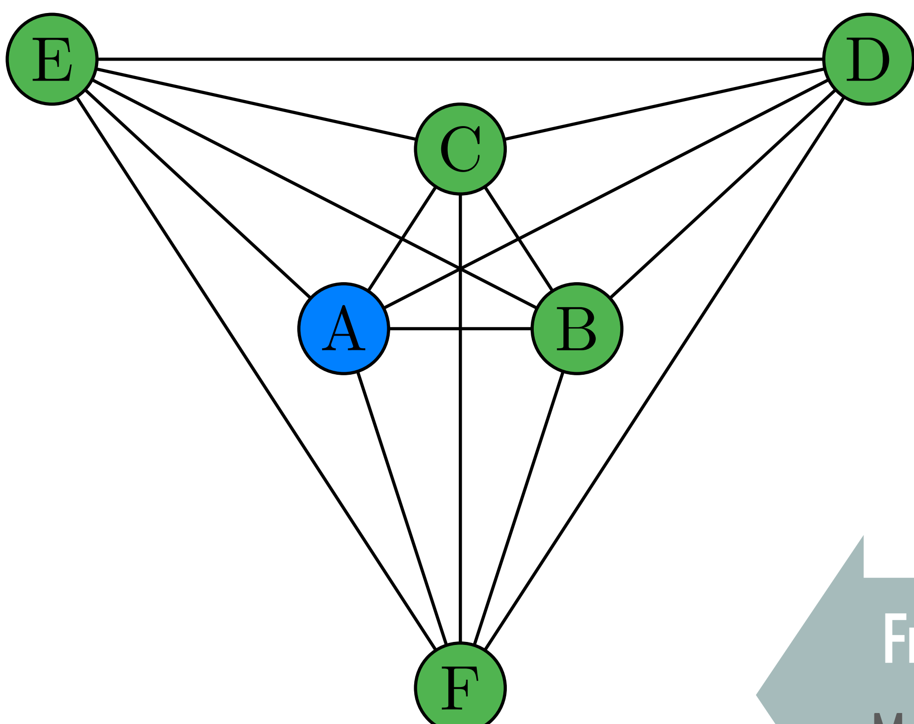


SIDE OBSERVATIONS – AN INTERMEDIATE GAME

Full-information

- observe losses of **all** actions
- example: Hedge

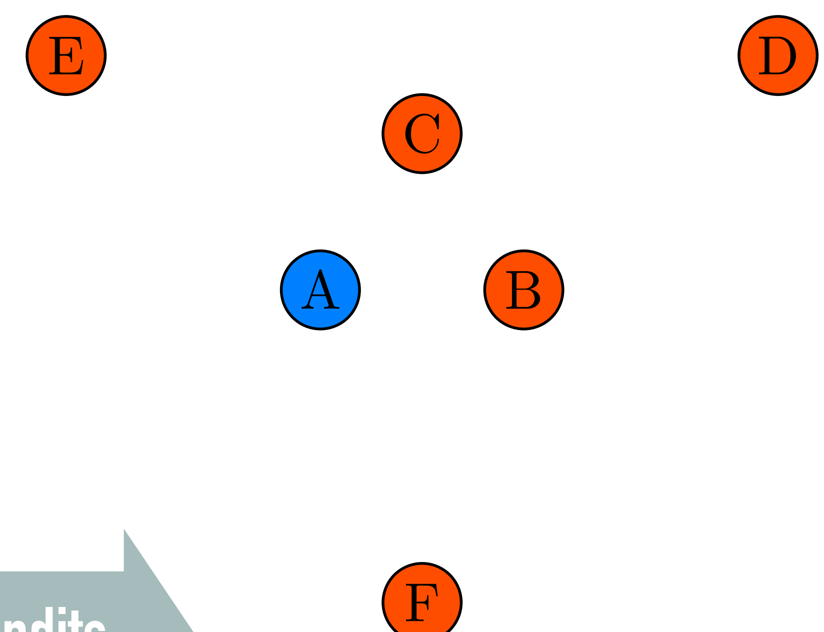
$$R_T = \widetilde{\mathcal{O}}(\sqrt{T})$$



Bandits

- observe losses of **the chosen** action
- example: EXP3

$$R_T = \widetilde{\mathcal{O}}(\sqrt{NT})$$

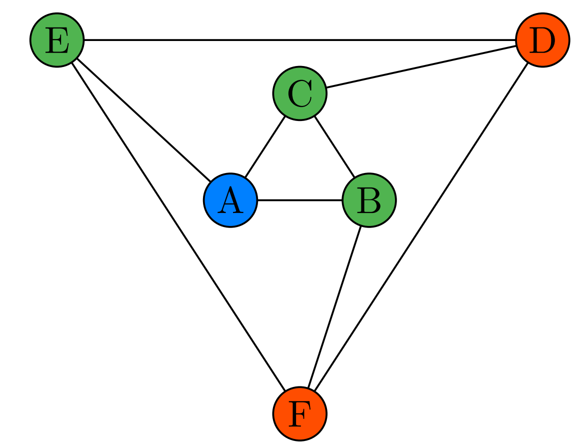


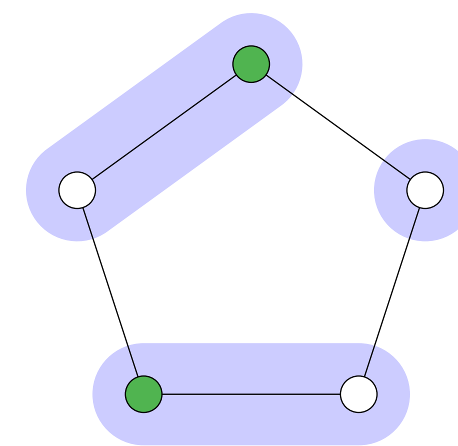
From Experts to Bandits

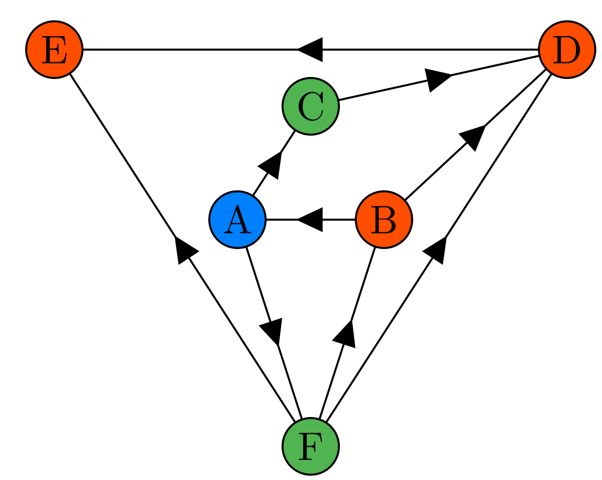
Mannor and Shamir 2011

KNOWLEDGE OF OBSERVATION GRAPHS

- ELP (Mannor and Shamir 2011)
  - **EXP3** - with **"LP balanced exploration"**
  - undirected $O(\sqrt{(\alpha T)})$ ✅ – needs to know G_t
  - directed case $O(\sqrt{(cT)})$ – needs to know G_t
- EXP3-SET (Alon, Cesa-Bianchi, Gentile, Mansour, 2013)
  - undirected $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
- EXP3-DOM (Alon, Cesa-Bianchi, Gentile, Mansour, 2013)
  - directed $O(\sqrt{(\alpha T)})$ ✅ – need to know G_t
  - **calculates dominating set**







Reminder: Exp3 algorithms in general

- **Compute weights** using loss estimates $\hat{\ell}_{t,i}$.

$$w_{t,i} = \exp\left(-\eta \sum_{s=1}^{t-1} \hat{\ell}_{s,i}\right)$$

- **Play action** I_t such that

$$\mathbb{P}(I_t = i) = p_{t,i} = \frac{w_{t,i}}{W_t} = \frac{w_{t,i}}{\sum_{j=1}^{N} w_{t,j}}$$

- **Update loss estimates** (using observability graph)

How the algorithms approach to bias variance tradeoff?

Bias variance tradeoff approaches

- Approach of Mixing
  - Bias sampling distribution $\mathbf{p}_t$ over actions
    - $\mathbf{p}'_t = (1 - \gamma)\mathbf{p}_t + \gamma \mathbf{s}_t$ – mixed distribution
    - $\mathbf{s}_t$ – probability distribution which supports exploration
  - Loss estimates $\hat{\ell}_{t,i}$ are unbiased
- Approach of Implicit eXploration (**IX**)
  - Bias loss estimates $\hat{\ell}_{t,i}$
    - Biased loss estimates $\implies$ biased weights
    - Biased weights $\implies$ biased probability distribution
  - No need for mixing

Is there a difference in a traditional non-graph case? Not much

Big difference in graph feedback case!

Algorithm 1 EXP3-IX

Input: Set of actions $\mathcal{S} = [d]$,
  parameters $\gamma_t \in (0,1)$, $\eta_t > 0$ for $t \in [T]$.
for $t = 1$ **to** T **do**
  $w_{t,i} \leftarrow (1/d) \exp\left(-\eta_t \widehat{L}_{t-1,i}\right)$ for $i \in [d]$
  An adversary privately chooses losses $\ell_{t,i}$ for $i \in [d]$ and generates a graph G_t
  $W_t \leftarrow \sum_{i=1}^{d} w_{t,i}$
  $p_{t,i} \leftarrow w_{t,i}/W_t$
  Choose $I_t \sim \boldsymbol{p}_t = (p_{t,1}, \ldots, p_{t,d})$
  Observe graph G_t
  Observe pairs $\{i, \ell_{t,i}\}$ for $(I_t \to i) \in G_t$
  $o_{t,i} \leftarrow \sum_{(j \to i) \in G_t} p_{t,j}$ for $i \in [d]$
  $\hat{\ell}_{t,i} \leftarrow \frac{\ell_{t,i}}{o_{t,i} + \gamma_t} \mathbb{1}_{\{(I_t \to i) \in G_t\}}$ for $i \in [d]$
end for

Benefits of the **implicit exploration**

- **no need to know the graph before**
- no need to estimate dominating set
- no need for doubling trick
- no need for aggregation

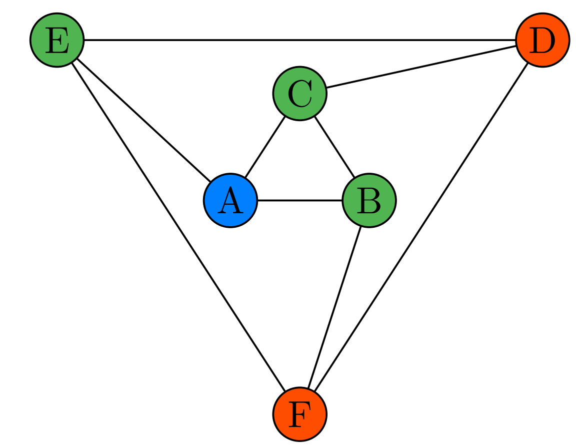
$$R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha T \ln N}\right)$$

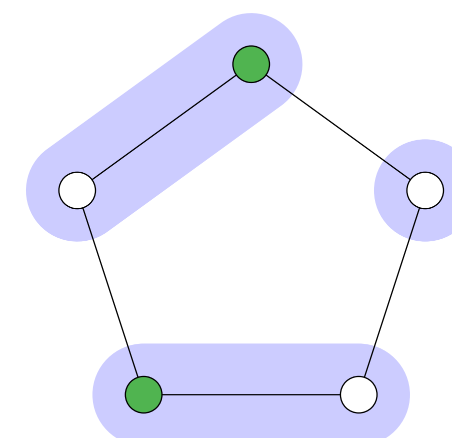
Optimistic bias for the loss estimates

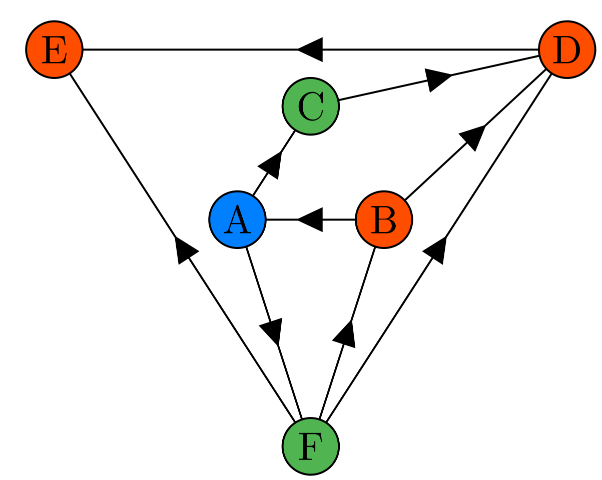
$$\mathbb{E}[\hat{\ell}_{t,i}] = \frac{\ell_{t,i}}{o_{t,i} + \gamma} o_{t,i} + 0(1 - o_{t,i}) = \ell_{t,i} - \ell_{t,i} \frac{\gamma}{o_{t,i} + \gamma} \leq \ell_{t,i}$$

FOLLOW UPS

- EXP3-IX (Kocák, Neu, MV, Munos, 2014)
  - directed $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
- EXP3.G (Alon, Cesa-Bianchi, Dekel, Koren, 2015)
  - directed $O(\sqrt{(\alpha T)})$ ✅ does not need to know G_t ✅
  - mixes uniform distribution
  - more general algorithm for settings **beyond bandits**
  - high-probability bound
- Neu 2015: high-probability bound for EXP3-IX

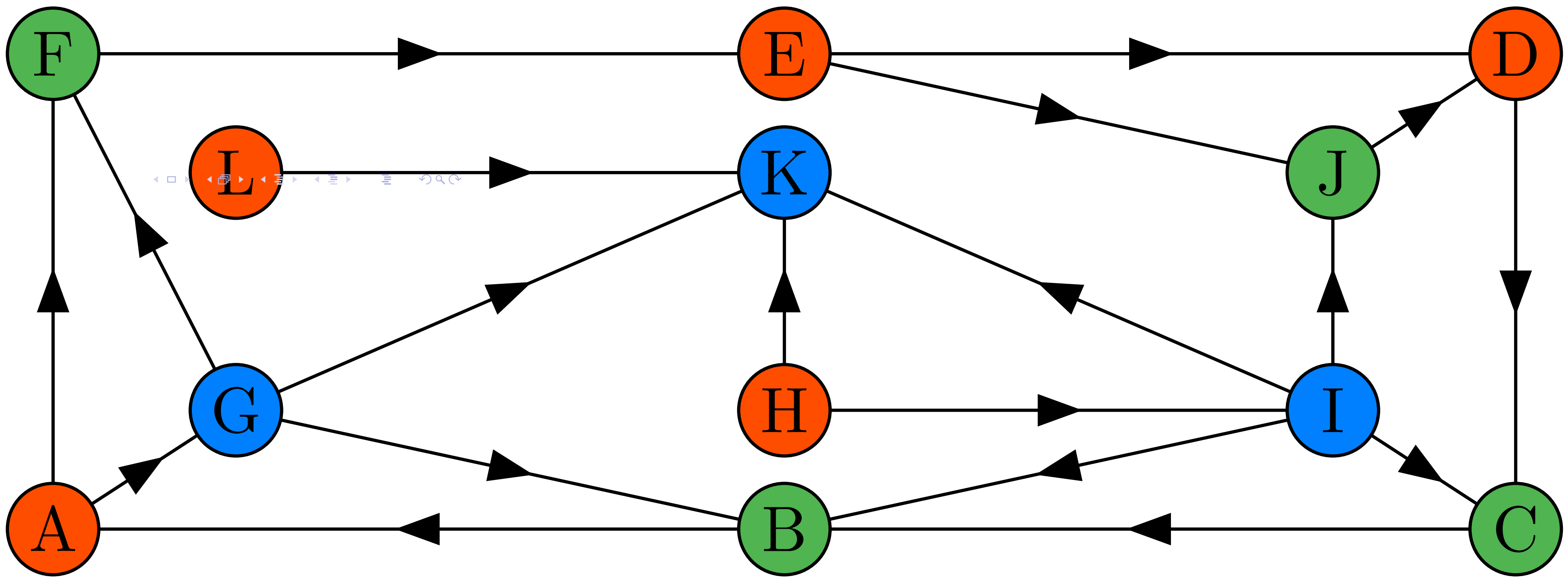






COMPLEX GRAPH ACTIONS

Example: online shortest path semi-bandits with observing traffic on the side streets



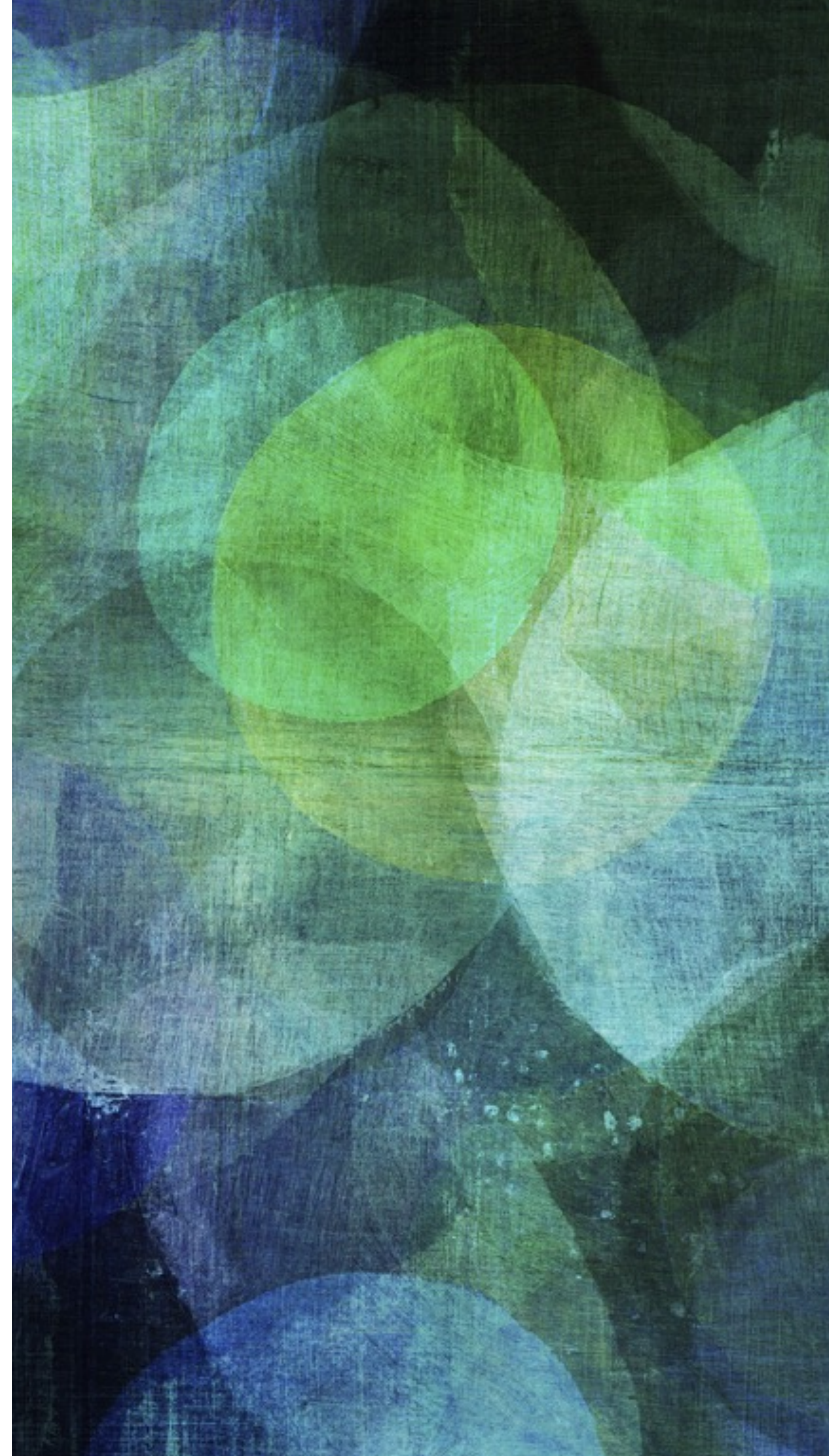
- Play action $\mathbf{V}_t \in S \subset \{0,1\}^N$, $\|\mathbf{v}\|_1 \leq m$ from all $\mathbf{v} \in S$
- Obtain losses $\mathbf{V}_t^\mathsf{T} \ell_t$
- Observe additional losses according to the graph

$$R_T = \widetilde{\mathcal{O}}\left(m^{3/2}\sqrt{\sum_{t=1}^{T}\alpha_t}\right) = \widetilde{\mathcal{O}}\left(m^{3/2}\sqrt{\overline{\alpha}T}\right)$$

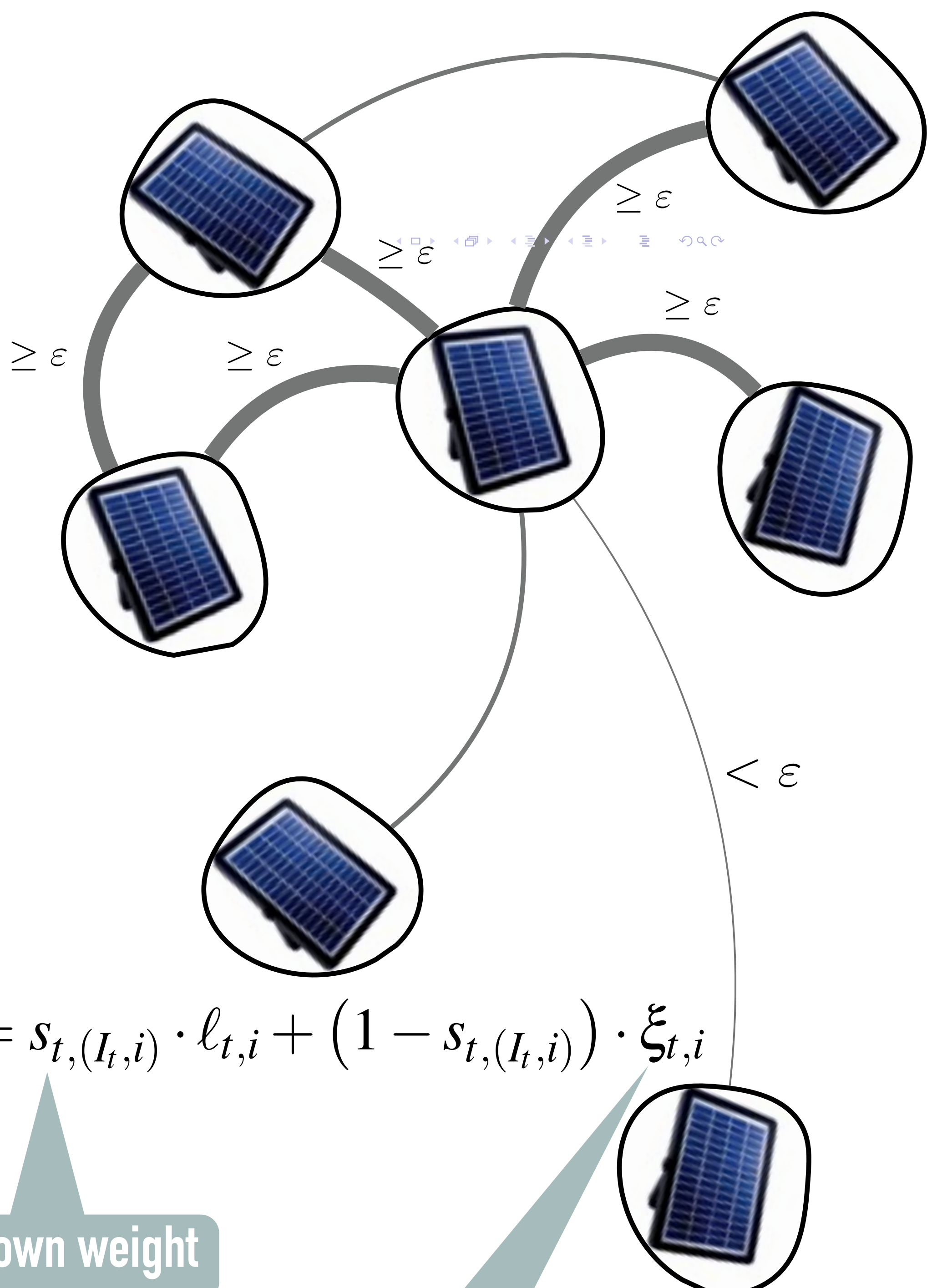
Kocák, Neu, MV: **Online learning with noisy side observations**, AISTATS 2016

GRAPH BANDITS WITH NOISY SIDE OBSERVATIONS

exploiting side observations that can be perturbed by certain level of noise



NOISY SIDE OBSERVATIONS



$$c_{t,i} = s_{t,(I_t,i)} \cdot \ell_{t,i} + \left(1 - s_{t,(I_t,i)}\right) \cdot \xi_{t,i}$$

known weight

zero-mean noise

Want: only reliable information!

1) If we know the perfect cutoff ε

- reliable: use as exact
- unreliable: rubbish

then we can improve over pure bandit setting!

2) Treating noisy observation induces bias

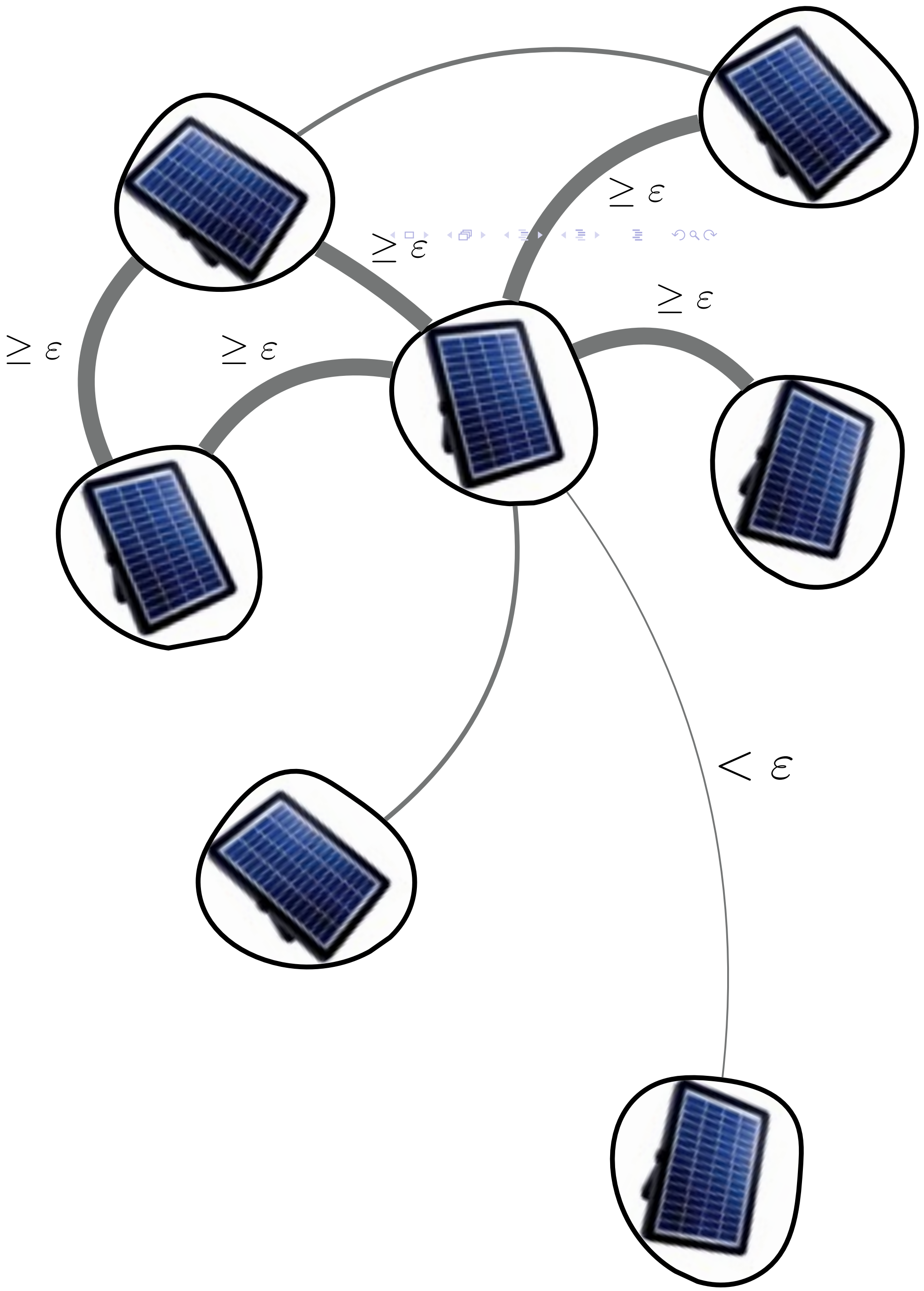
What can we hope for?

$$\tilde{\mathcal{O}}\left(\sqrt{1T}\right) \leq \qquad \leq \tilde{\mathcal{O}}\left(\sqrt{NT}\right)$$

effective independence number

Can we learn without knowing either ε or α^* ?

NOISY SIDE OBSERVATIONS



Threshold estimate $R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha^\star T}\right)$

$$\widehat{\ell}^{(\mathrm{T})}_{t,i} = \frac{c_{t,i}\mathbb{I}_{\{s_{t,(I_t,i)} \geq \varepsilon_t\}}}{\sum_{j=1}^{N} p_{t,j} s_{t,(j,i)} \mathbb{I}_{\{s_{t,(j,i)} \geq \varepsilon_t\}} + \gamma_t}$$

WIX estimate $R_T = \widetilde{\mathcal{O}}\left(\sqrt{\alpha^\star T}\right)$

$$\widehat{\ell}_{t,i} = \frac{s_{t,(I_t,i)} \cdot c_{t,i}}{\sum_{j=1}^{N} p_{t,j} s^2_{t,(j,i)} + \gamma_t}$$

Since $\alpha^* \leq \alpha(1)/1 \leq N$

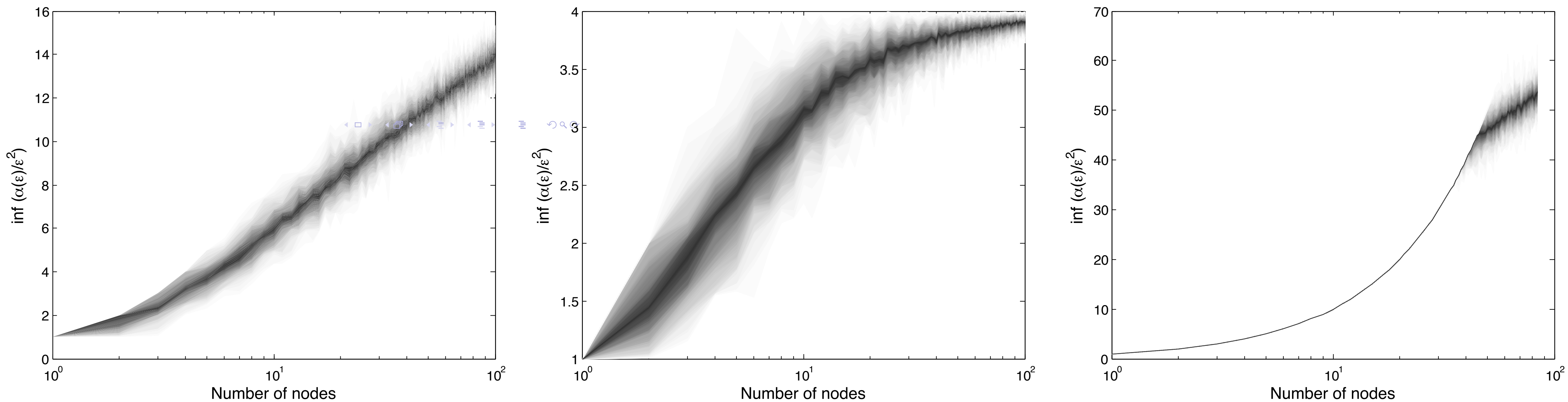
incorporating noisy observations does not hurt

$$\widetilde{\mathcal{O}}\left(\sqrt{\alpha^\star T}\right) \leq \widetilde{\mathcal{O}}\left(\sqrt{N T}\right)$$

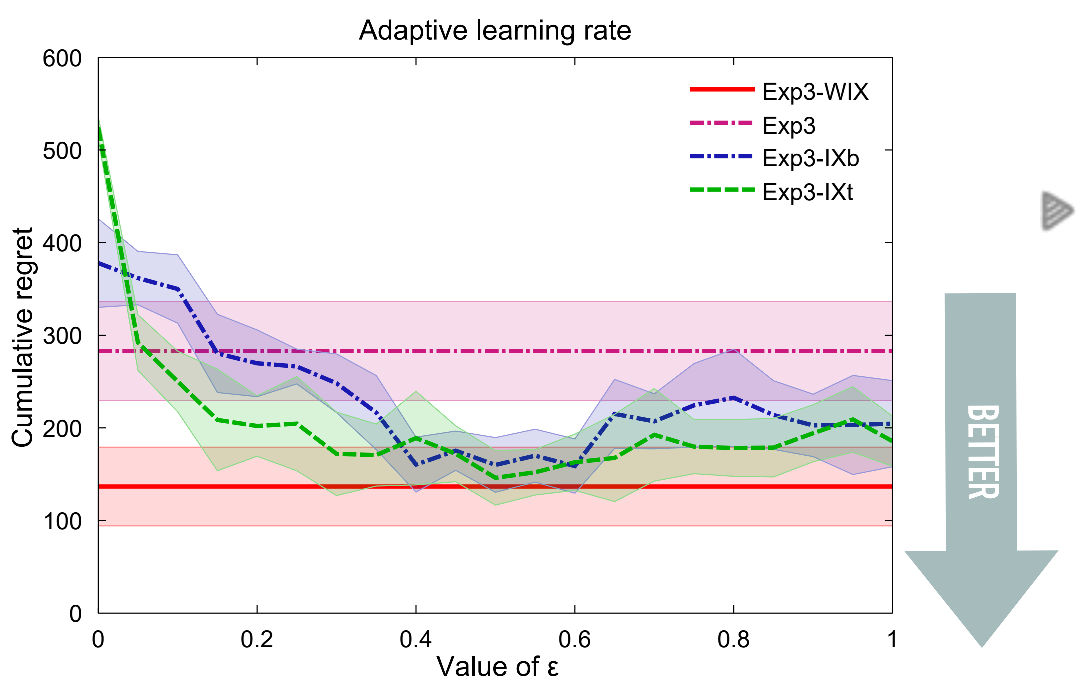
But how much does it help?

EMPIRICAL RESULTS

EMPIRICAL α^* FOR RANDOM GRAPHS WITH IID WEIGHTS



(a) $U(0,1)$ weights (b) $U(\frac{1}{2},1)$ weights (c) $U(0,\frac{1}{2})$ weights



- **special case:** if s_{ij} is either 0 or ε than $\alpha^* = \alpha/\varepsilon^2$
  - For this special case, there is a matches $\Theta(\sqrt{(\alpha T)}/\varepsilon)$ by Wu, György, Szepesvári, 2015.

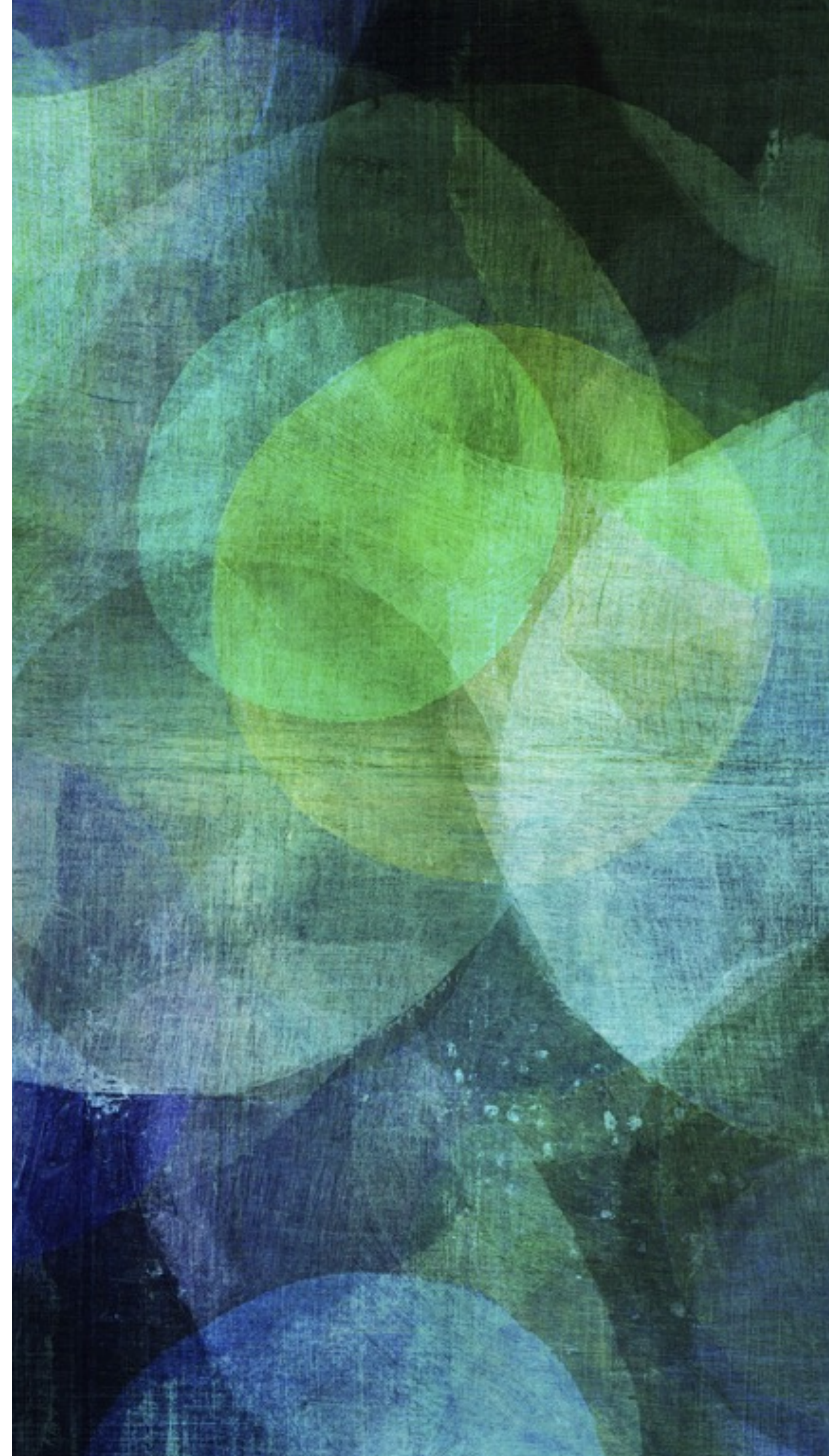
NEW DIRECTIONS: UNKNOWN GRAPHS!

- **Learning on the graph while learning the graph?**
  - **most of algorithms require (some) knowledge of the graph**
  - not always available to the learner
- Question: Can we learn faster without knowing the graphs?
  - example: social network provider has little incentive to reveal the graphs to advertisers
- Answer: **Cohen, Hazan, and Koren:** Online learning with **feedback** graphs without the graphs (ICML June 19-24, 2016)
  - **NO!** (in general we cannot, but possible in the stochastic case)
- Coming up next:
  - **Erdös-Rényi side observation graphs** (UAI June 25-26, 2016)

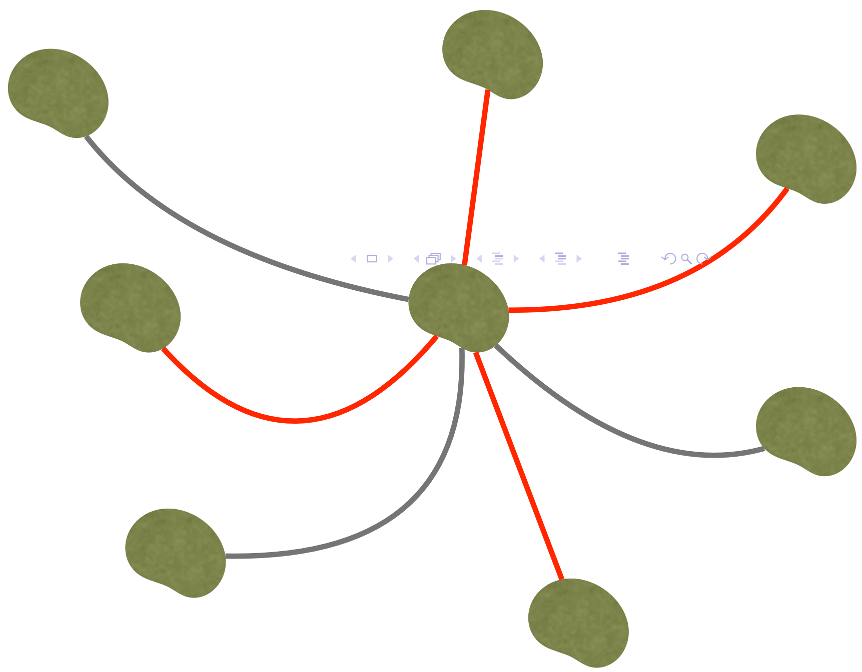
Kocák, Neu, MV: **Online learning with Erdos-Rényi side-observation graphs**
UAI 2016

GRAPH BANDITS WITH ERDÖS-RÉNYI OBSERVATIONS

side observations from graph generators



PROTOCOL FOR ERDÖS-RÉNYI GRAPHS



Every round **t** the learner

- picks a node I_t
- suffers loss for I_t
- receives feedback
  - for I_t
  - for every other node with probability r_t

is loss of i observed?

true loss

$$\widehat{\ell}^{\star}_{t,i} = \frac{O_{t,i}\ell_{t,i}}{p_{t,i} + (1 - p_{t,i})r_t}$$

probability of picking i

probability of side observation

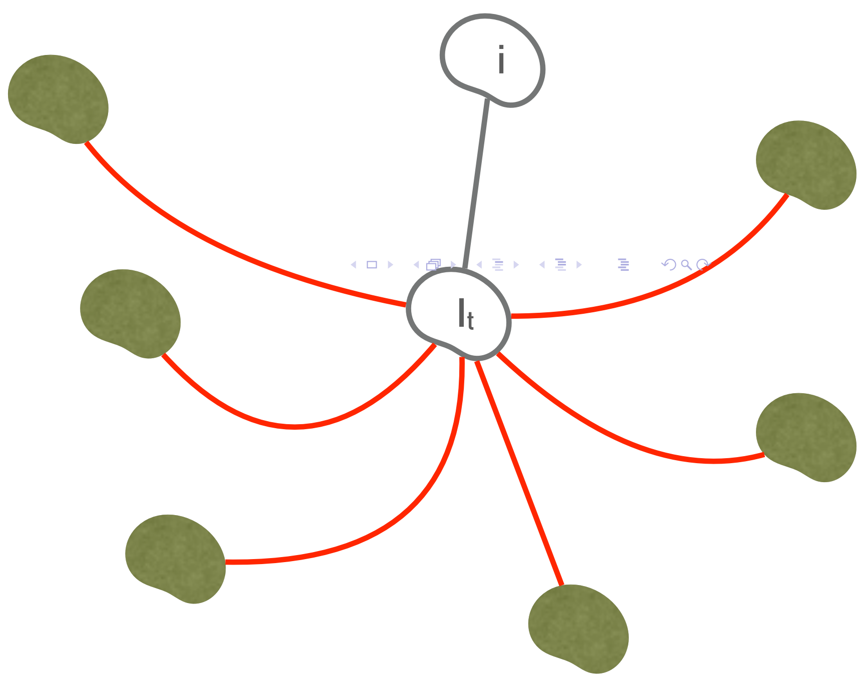
Regret of Exp3-SET (Alon et al. 2013):

$$\mathcal{O}\left(\sqrt{\textstyle\sum_t (1/r_t)(1-(1-r_t)^N)\log N}\right)$$

How to estimate r_t in every round when it is changing?

How to estimate losses without the knowledge of r_t?

PROTOCOL FOR ERDÖS-RÉNYI GRAPHS



- N-2 samples from Bernoulli(r_t) ... R(k)
- N-2 samples from p_{ti} ... P(k)
- O'(k) = P(k) + (1-P(k))R(k)
- G_{ti} = min{k : O'(k) = 1} U {N-1}
- $E[G_{ti}] \approx 1/(p_{ti} + (1-p_{ti})r_t)$

$$\widehat{\ell}_{t,i} = G_{t,i} O_{t,i} \ell_{t,i}$$

is loss of i observed?

true loss

$$\widehat{\ell}^{\star}_{t,i} = \frac{O_{t,i}\ell_{t,i}}{p_{t,i} + (1 - p_{t,i})r_t}$$

probability of picking i

probability of side observation

If $r_t \geq (\log T)/(2N-2)$ then

$$\mathcal{O}\left(\sqrt{\log N \sum_{t=1}^{T} \frac{1}{r_t}}\right)$$

Lower bound (Alon et al. 2013) $\Omega(\sqrt{T/r})$

Get rid of $r_t \geq (\log T)/(2N-2)$?

MORE GRAPH BANDITS AND BEYOND!

Noga Alon et al. (2015) Beyond bandits. Complete characterization: Bártok et al. (2014)

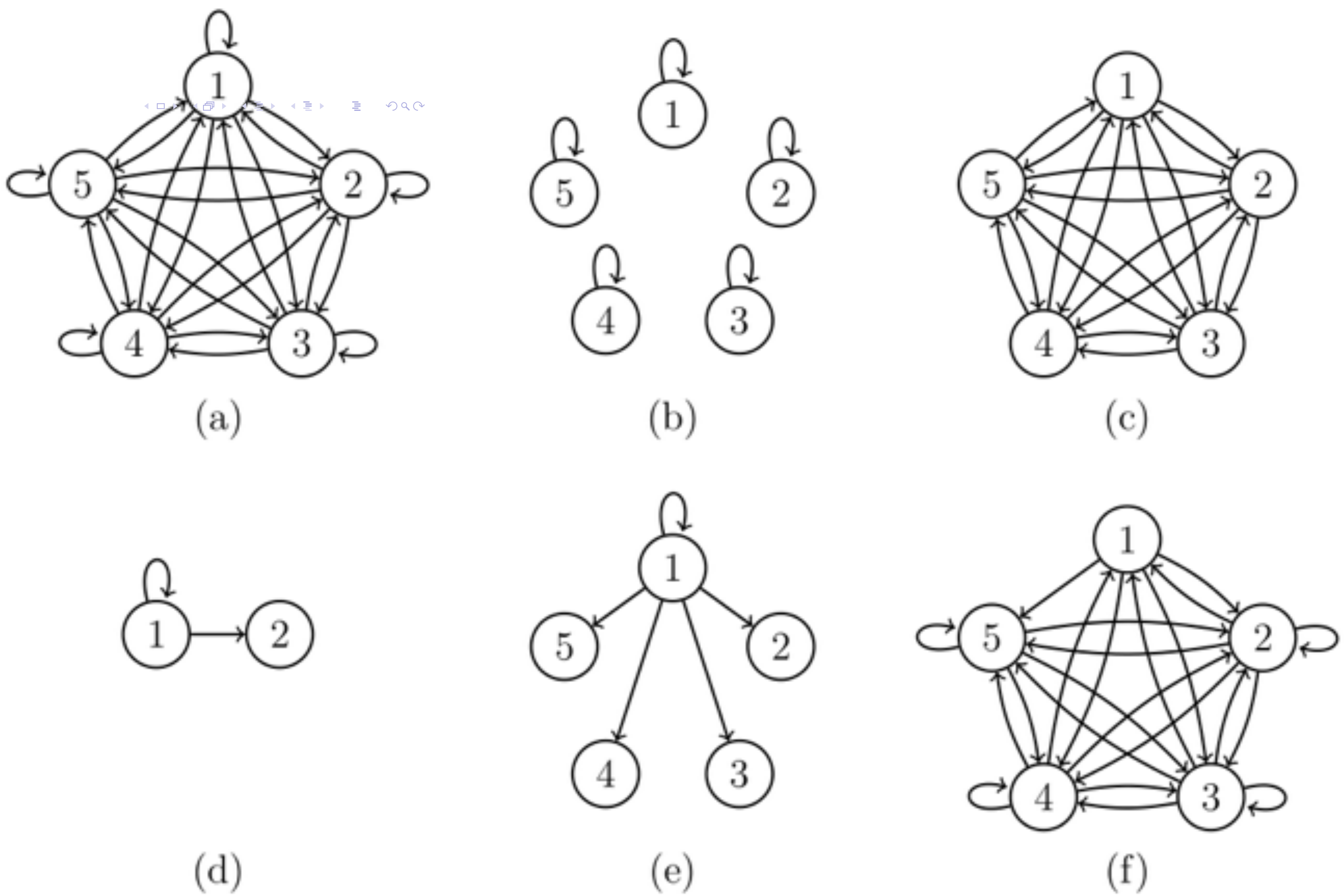
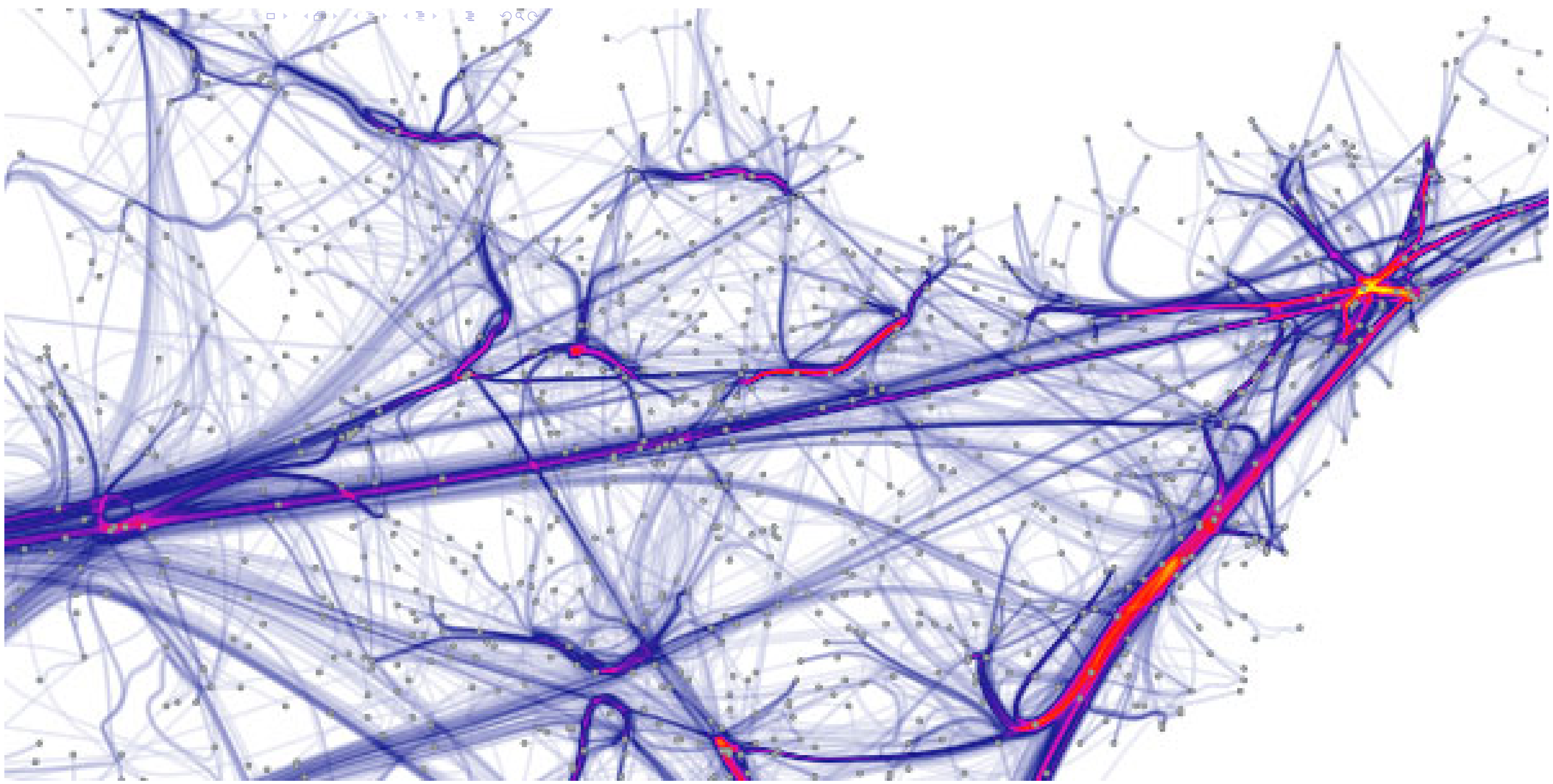


Figure 1: Examples of feedback graphs: (a) *full feedback*, (b) *bandit feedback*, (c) *loopless clique*, (d) *apple tasting*, (e) *revealing action*, (f) a clique minus a self-loop and another edge.

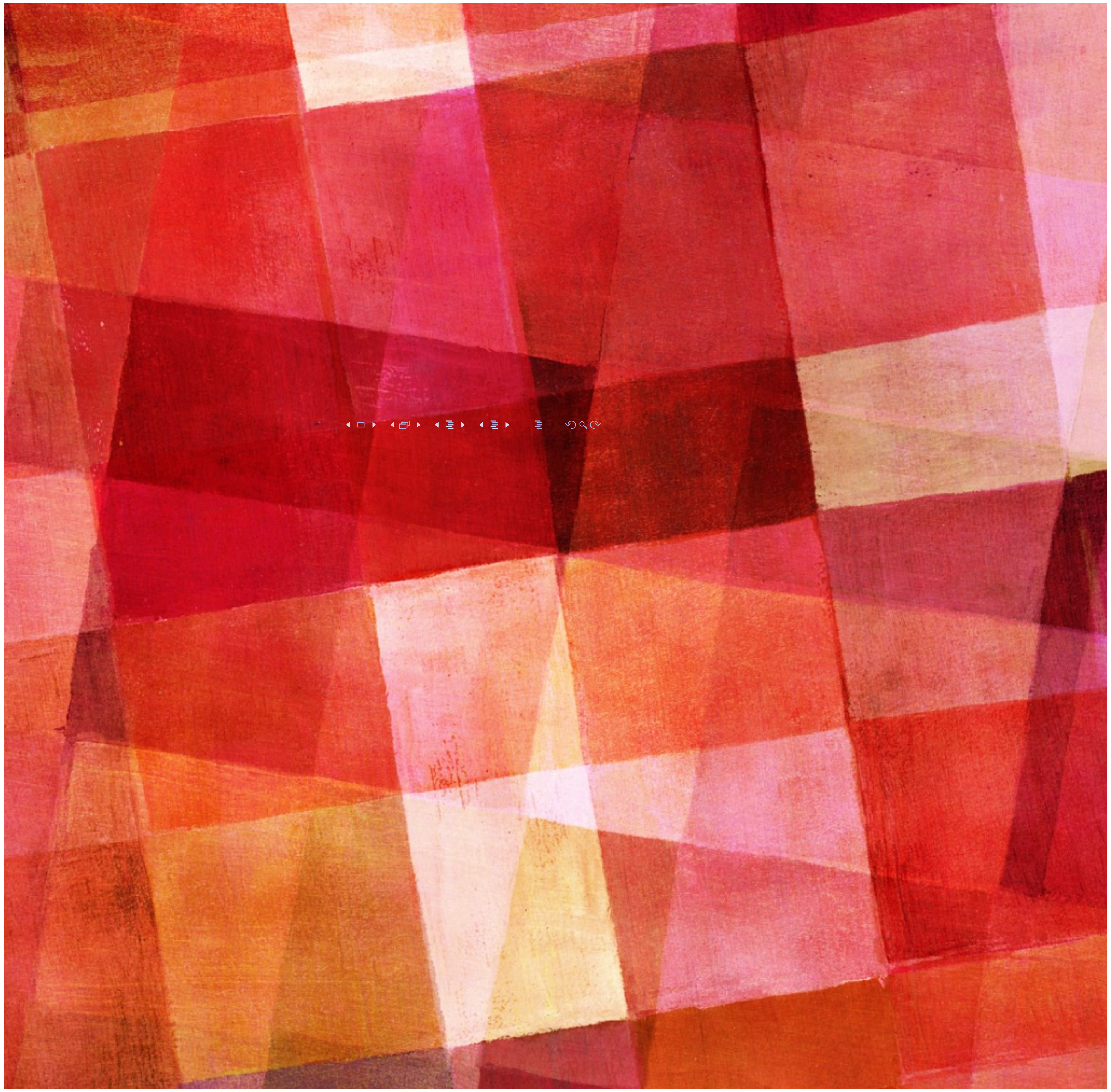
LAST WORDS …

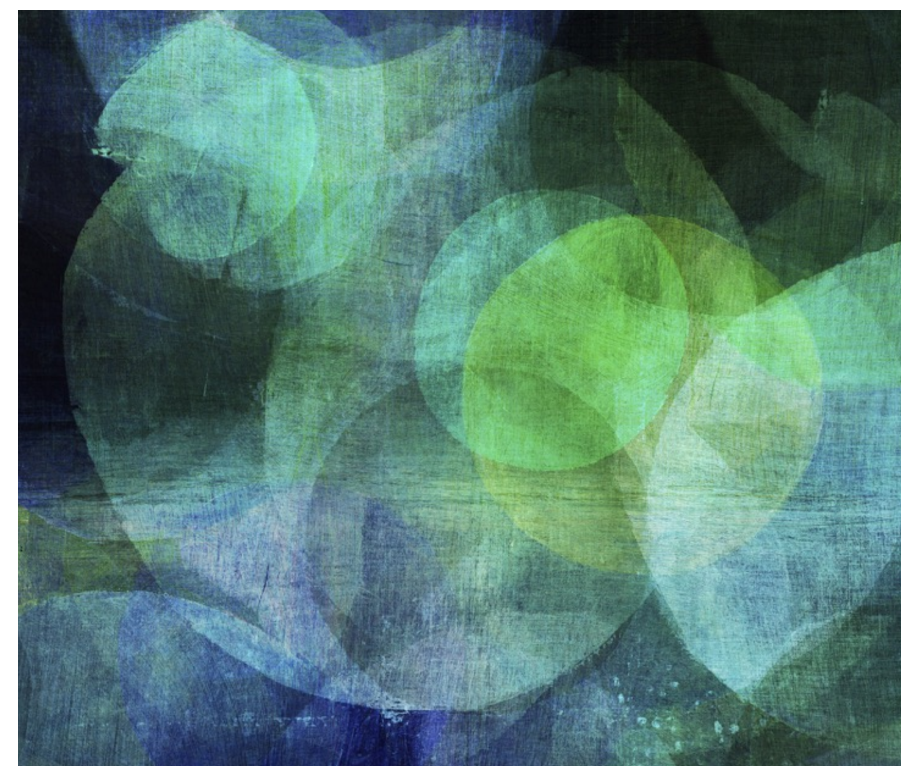
Survey: http://researchers.lille.inria.fr/~valko/hp/publications/valko2016bandits.pdf **(Part I)**

1) good luck with the projects 2) AlteGrad follows this course 3) see you at projects talks



THAT'S ALL – THANK YOU!





Michal Valko, SequeL, Inria Lille - Nord Europe, michal.valko@inria.fr
http://researchers.lille.inria.fr/~valko/hp/