# Adaptive Rumor Spreading

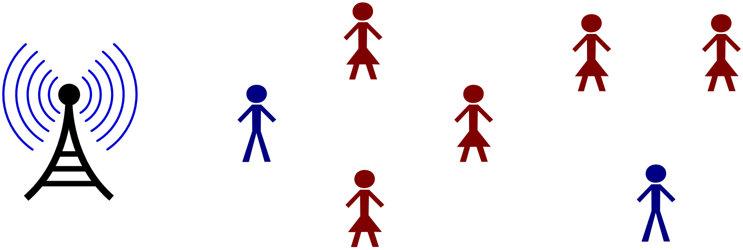
José Correa [1] Marcos Kiwi [1]
Neil Olver [2] Alberto Vera [1]

[1]Universidad de Chile

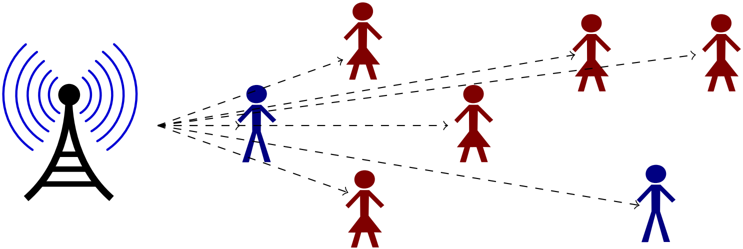
[2]VU Amsterdam and CWI

July 27, 2015

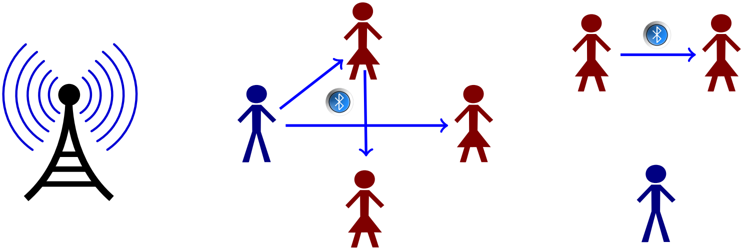
# The situation



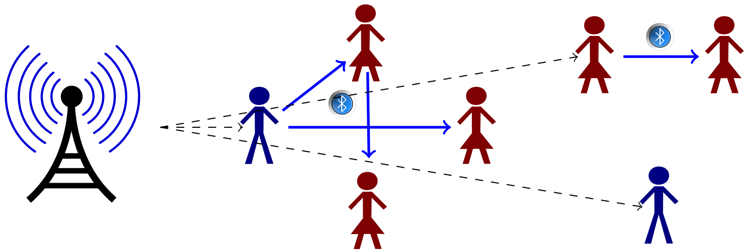
# The situation



# The situation



# The situation



# Introduction

- Rumors in social networks: contents, updates, new technology, etc.

# Introduction

- Rumors in social networks: contents, updates, new technology, etc.
- In viral marketing campaigns, the selection of vertices is crucial. Domingos and Richardson (2001)

# Introduction

- Rumors in social networks: contents, updates, new technology, etc.
- In viral marketing campaigns, the selection of vertices is crucial. Domingos and Richardson (2001)
- An agent (service provider) wants to efficiently speed up the communication process.

# Rumor spreading

- Models differ in time and communication protocol. Demers et al. (1987) and Boyd et al. (2006)

# Rumor spreading

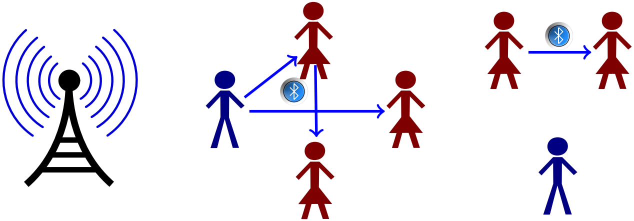
- Models differ in time and communication protocol. Demers et al. (1987) and Boyd et al. (2006)
- In simple cases, the time to activate all the network is mostly understood.

# Rumor spreading

- Models differ in time and communication protocol. Demers et al. (1987) and Boyd et al. (2006)
- In simple cases, the time to activate all the network is mostly understood.
- Even in random networks the estimates are logarithmic in the number of nodes. Doerr et al. (2012) and Chierichetti et al. (2011)

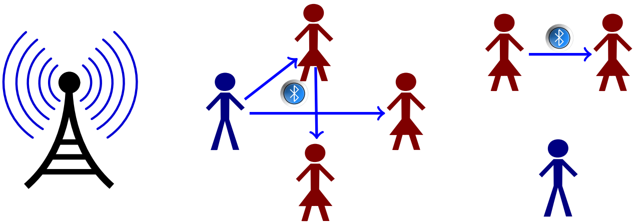
# Opportunistic networks

- We have an overload problem, an option is to exploit opportunistic communications.



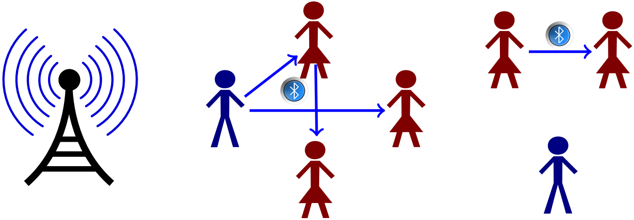
# Opportunistic networks

- We have an overload problem, an option is to exploit opportunistic communications.
- A fixed deadline scenario has been studied heuristically along with real large-scale data. Whitbeck et al. (2011)



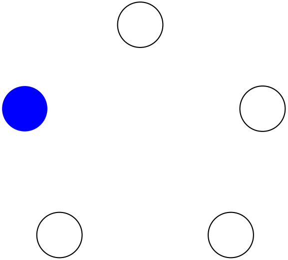
# Opportunistic networks

- We have an overload problem, an option is to exploit opportunistic communications.
- A fixed deadline scenario has been studied heuristically along with real large-scale data. Whitbeck et al. (2011)
- Control theory based algorithms greatly outperform static ones. Sciancalepore et al. (2014)



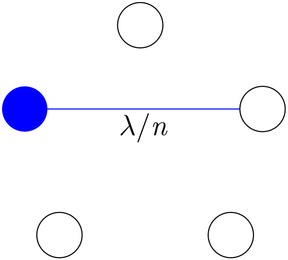
# The model

- Bob communicates and shares information.



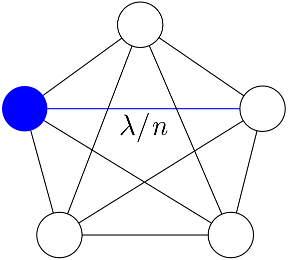
# The model

- Bob communicates and shares information.
- Bob meets Alice according to a Poisson process of rate $\lambda/n$.



# The model

- Bob communicates and shares information.
- Bob meets Alice according to a Poisson process of rate $\lambda/n$.
- Every pair of nodes can meet and gossip.



# The problem

- There is a unit cost for pushing the rumor.
- Opportunistic communications have no cost.
- At time $\tau$ all of the graph must be active.

# The problem

- There is a unit cost for pushing the rumor.
- Opportunistic communications have no cost.
- At time $\tau$ all of the graph must be active.

We want a strategy that minimizes the overall number of pushes.

# Adaptive and non-adaptive

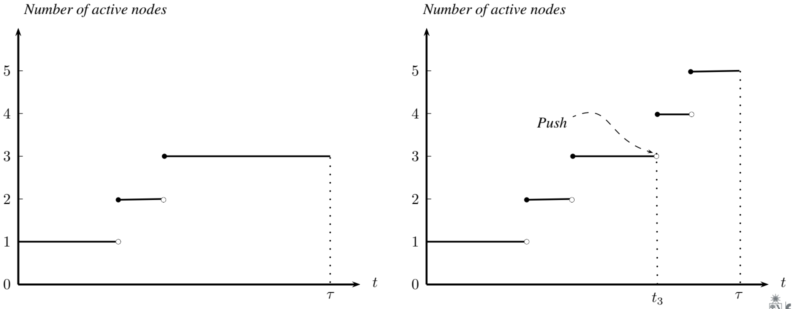
- A non-adaptive strategy pushes only at times $t = 0$ and $t = \tau$.

# Adaptive and non-adaptive

- A non-adaptive strategy pushes only at times $t = 0$ and $t = \tau$.
- An adaptive strategy may push at any time, with the full knowledge of the process' evolution.

# Adaptive and non-adaptive

- A non-adaptive strategy pushes only at times $t = 0$ and $t = \tau$.
- An adaptive strategy may push at any time, with the full knowledge of the process' evolution.



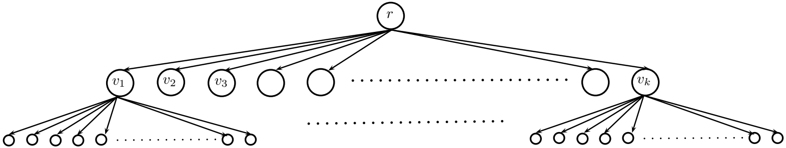
# Main result

Define the adaptivity gap as the ratio between the expected costs of non-adaptive and adaptive.

## Theorem

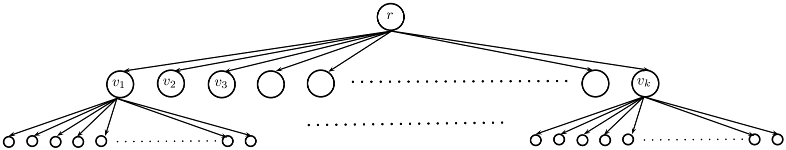
*In the complete graph the adaptivity gap is constant.*

# Adaptive can be arbitrarily better



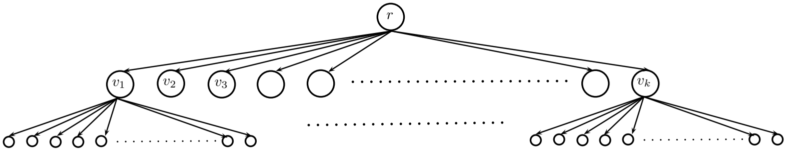
# Adaptive can be arbitrarily better

- With a small deadline, non-adaptive activates all of the $v_i$'s.
- Adaptive activates only the root, then at some time $t'$ pushes to the inactive $v_i$'s.

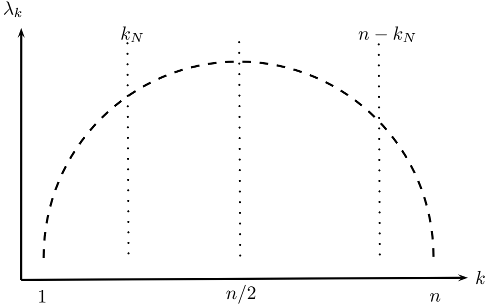


# Adaptive can be arbitrarily better

- With a small deadline, non-adaptive activates all of the $v_i$'s.
- Adaptive activates only the root, then at some time $t'$ pushes to the inactive $v_i$'s.
- An adaptivity gap of $\frac{\log k}{\log \log k}$ is easy to prove.



# Non-adaptive

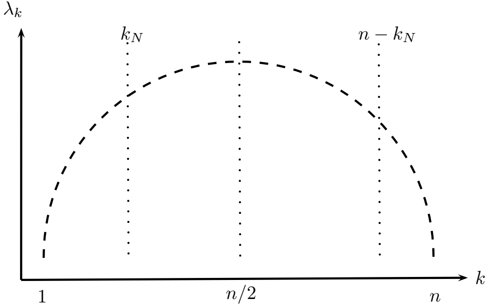



$\lambda = 1$.

$\lambda_k := \frac{k(n-k)}{n}$.

# Non-adaptive

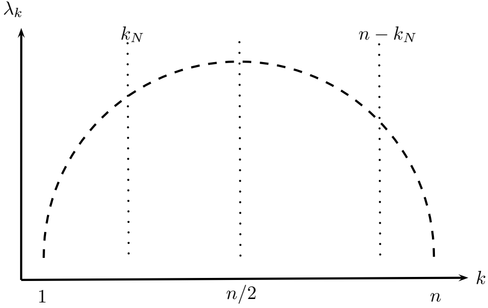
- Optimal non-adaptive pays almost the same at $t = 0$ and at $t = \tau$.




$\lambda = 1$.
$\lambda_k := \frac{k(n-k)}{n}$.

# Non-adaptive

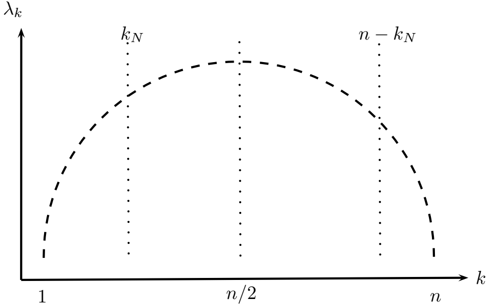
- Optimal non-adaptive pays almost the same at $t = 0$ and at $t = \tau$.
  - A 2-approximation is easy to see.




$\lambda = 1$.
$\lambda_k := \frac{k(n-k)}{n}$.

# Non-adaptive

- Optimal non-adaptive pays almost the same at $t = 0$ and at $t = \tau$.
  - A 2-approximation is easy to see.
- Non-adaptive does not push more than $n/2$ rumors. Therefore, neither adaptive.



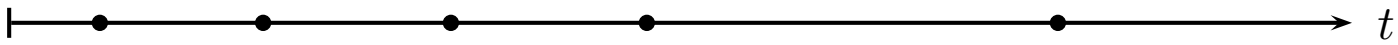

$\lambda = 1.$
$\lambda_k := \frac{k(n-k)}{n}.$

# Big deadline: $\tau \geq (2+\delta)\log n$

- Starting from a single active node, the time until everyone is active is $2\log n + \mathcal{O}(1)$.
- The time is exponentially concentrated. Jason (1999)
- Just starting with one node has cost $1+\varepsilon$, therefore adaptivity does not help.

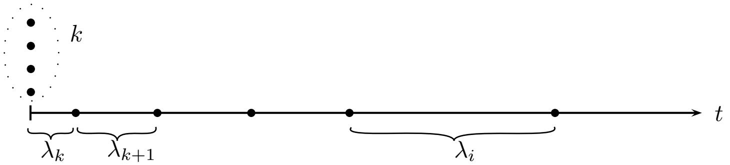
# Small deadline: $\tau \leq 2 \log \log n$

- A Poisson process of unit rate gives the randomness.



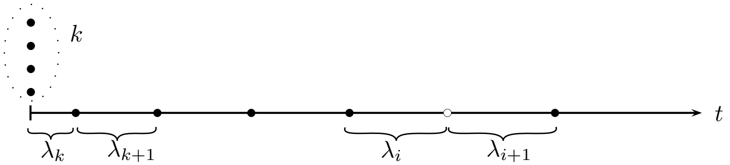
# Small deadline: $\tau \leq 2 \log \log n$

- A Poisson process of unit rate gives the randomness.
- Given the points $S_i$ and $S_{i+1}$, the rescaling $\frac{S_{i+1}-S_i}{\lambda_i}$ is the inter-arrival time.



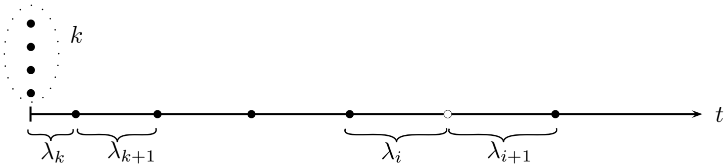
# Small deadline: $\tau \leq 2 \log \log n$

- A Poisson process of unit rate gives the randomness.
- Given the points $S_i$ and $S_{i+1}$, the rescaling $\frac{S_{i+1}-S_i}{\lambda_i}$ is the inter-arrival time.
- A push can be seen as adding a point.



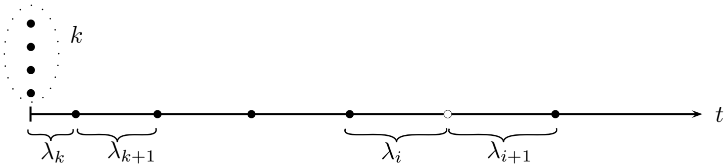
# Small deadline (cont.): $\tau \leq 2 \log \log n$

- A clairvoyant strategy knows the realization, therefore outperforms adaptive.



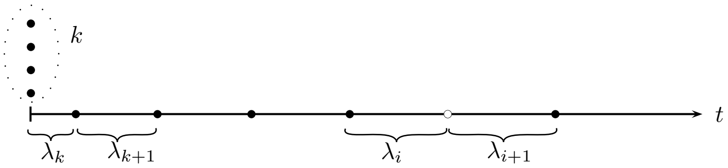
# Small deadline (cont.): $\tau \leq 2 \log \log n$

- A clairvoyant strategy knows the realization, therefore outperforms adaptive.
- We show that clairvoyant adds points only at the beginning.



# Small deadline (cont.): $\tau \leq 2 \log \log n$

- A clairvoyant strategy knows the realization, therefore outperforms adaptive.
- We show that clairvoyant adds points only at the beginning.
- Clairvoyant chooses the best number of initial pushes, given the realization.



# Small deadline (cont.): $\tau \leq 2 \log \log n$

Say we start with $k$ initial pushes.

# Small deadline (cont.): $\tau \leq 2 \log \log n$

Say we start with $k$ initial pushes.

- We know the inter-arrival distributions.
- We know the non-adaptive cost; it pays $\Omega(\frac{n}{\log n})$.

# Small deadline (cont.): $\tau \leq 2 \log \log n$

Say we start with $k$ initial pushes.

- We know the inter-arrival distributions.
- We know the non-adaptive cost; it pays $\Omega(\frac{n}{\log n})$.

## Lemma

*Clairvoyant is considerably better than non-adaptive with probability at most* $\frac{1}{n^2}$ .

# Small deadline (cont.): $\tau \leq 2 \log \log n$

Say we start with $k$ initial pushes.

- We know the inter-arrival distributions.
- We know the non-adaptive cost; it pays $\Omega(\frac{n}{\log n})$.

## Lemma

*Clairvoyant is considerably better than non-adaptive with probability at most* $\frac{1}{n^2}$ .

In this case we can prove the gap to be $1 + o(1)$.

# Other deadlines

**Insight:** adaptive interferes when the cost of pushing is less than or equal to that of not pushing, i.e.,

$$1 + \text{cost}(k + 1 \text{ active nodes}) \leq \text{cost}(k \text{ active nodes}). \quad (\star)$$

# Other deadlines

**Insight:** adaptive interferes when the cost of pushing is less than or equal to that of not pushing, i.e.,

$$1 + \text{cost}(k + 1 \text{ active nodes}) \leq \text{cost}(k \text{ active nodes}). \quad (\star)$$

The expected cost remains the same, it is a martingale, thus the condition should be met only a few times.

# Other deadlines

**Insight:** adaptive interferes when the cost of pushing is less than or equal to that of not pushing, i.e.,

$$1 + \text{cost}(k + 1 \text{ active nodes}) \leq \text{cost}(k \text{ active nodes}). \quad (\star)$$

The expected cost remains the same, it is a martingale, thus the condition should be met only a few times.

- A relaxed strategy pushes for free, but with certain conditions.

# Other deadlines

**Insight:** adaptive interferes when the cost of pushing is less than or equal to that of not pushing, i.e.,

$$1 + \text{cost}(k + 1 \text{ active nodes}) \leq \text{cost}(k \text{ active nodes}). \quad (\star)$$

The expected cost remains the same, it is a martingale, thus the condition should be met only a few times.

- A relaxed strategy pushes for free, but with certain conditions.
  - Pushes only when $(\star)$ holds.
  - Does not push after $n/2$.

# Other deadlines

**Insight:** adaptive interferes when the cost of pushing is less than or equal to that of not pushing, i.e.,

$$1 + \text{cost}(k + 1 \text{ active nodes}) \leq \text{cost}(k \text{ active nodes}). \quad (\star)$$

The expected cost remains the same, it is a martingale, thus the condition should be met only a few times.

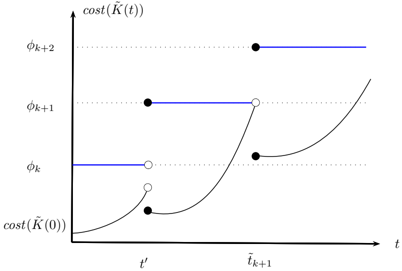
- A relaxed strategy pushes for free, but with certain conditions.
  - Pushes only when $(\star)$ holds.
  - Does not push after $n/2$.
- Relaxed outperforms adaptive.

# Other deadlines (cont.)

- Relaxed adaptive can be described by thresholds $\phi_k$.

# Other deadlines (cont.)

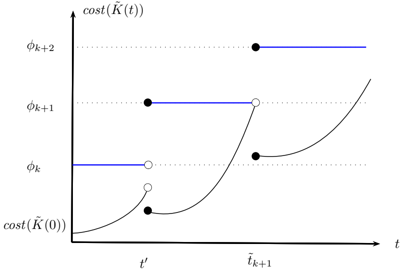
- Relaxed adaptive can be described by thresholds $\phi_k$.
- Let $\tilde{K}(t)$ be the number of active nodes at time $t$.



# Other deadlines (cont.)

- Relaxed adaptive can be described by thresholds $\phi_k$.
- Let $\tilde{K}(t)$ be the number of active nodes at time $t$.
- We transform the process:

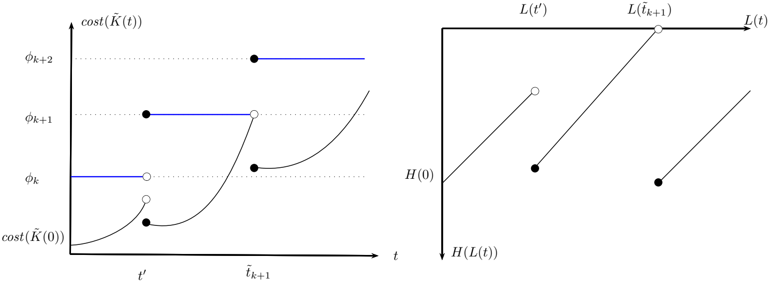
$$H(L(t)) := \lambda_{\tilde{K}(t)} \log \frac{\text{cost}(\tilde{K}(t))}{\phi_{\tilde{K}(t)}}.$$



# Other deadlines (cont.)

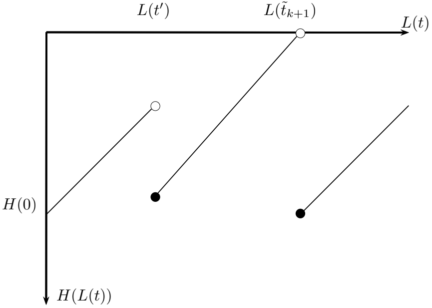
- Relaxed adaptive can be described by thresholds $\phi_k$.
- Let $\tilde{K}(t)$ be the number of active nodes at time $t$.
- We transform the process:

$$H(L(t)) := \lambda_{\tilde{K}(t)} \log \frac{\text{cost}(\tilde{K}(t))}{\phi_{\tilde{K}(t)}}.$$



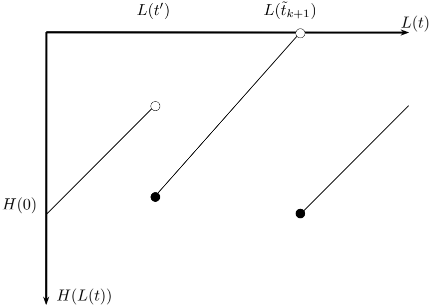
# Other deadlines (cont.)

- We show that each time $H$ touches zero, relaxed wins exactly 1 compared to non-adaptive.



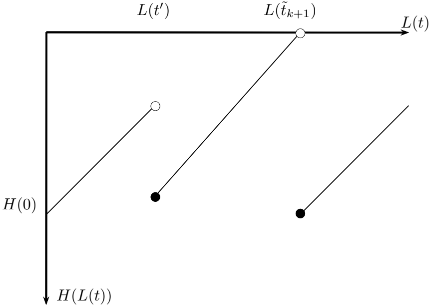
# Other deadlines (cont.)

- We show that each time $H$ touches zero, relaxed wins exactly 1 compared to non-adaptive.
- Essentially $H(s)$ is dominated by $s - 2\,\text{Poiss}(s)$.



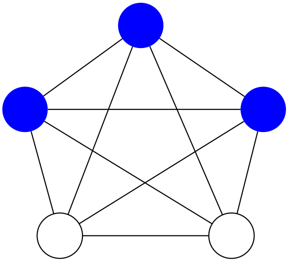
# Other deadlines (cont.)

- We show that each time $H$ touches zero, relaxed wins exactly 1 compared to non-adaptive.
- Essentially $H(s)$ is dominated by $s - 2\,\mathrm{Poiss}(s)$.
- The number of times $H(s)$ touches zero is constant.



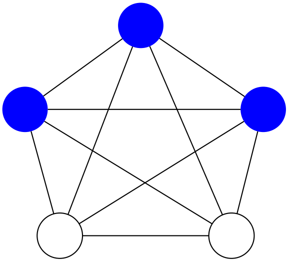
# Additional results

- The target set version has a constant adaptivity gap.

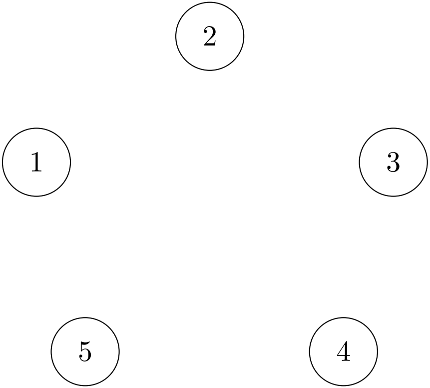


# Additional results

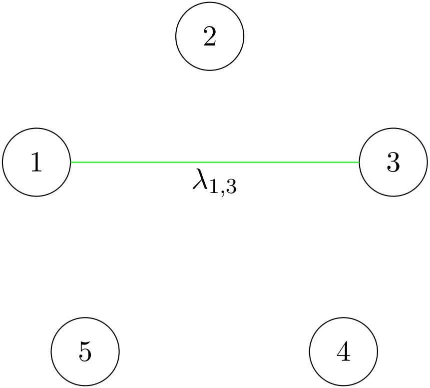
- The target set version has a constant adaptivity gap.
- The maximization problem has a $1 + o(1)$ adaptivity gap.



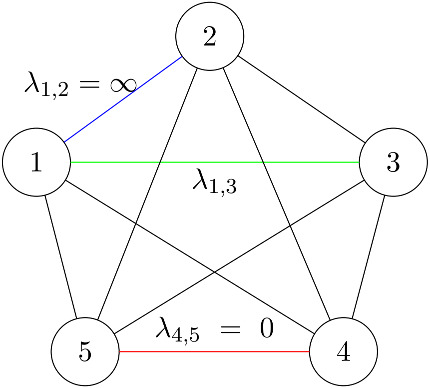
# General model



# General model

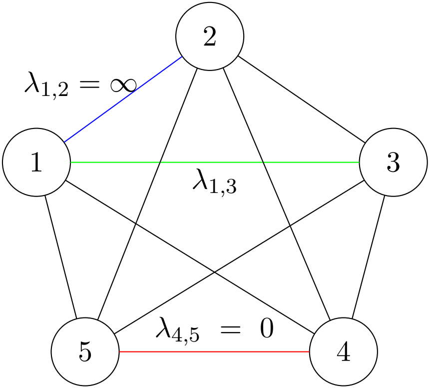


# General model



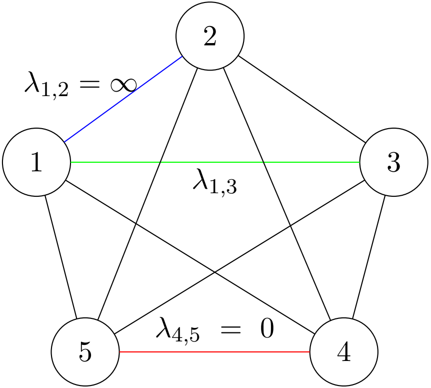
# General model

We need to keep track of the set of active nodes.



# General model

We need to keep track of the set of active nodes.
Even the non-adaptive problem is difficult in this setting!



# Conjectures and open problems

- Is there a broader class of graphs maintaining the constant gap result?
  - High conductance/connectivity.
  - Metric induced rates.

# Conjectures and open problems

- Is there a broader class of graphs maintaining the constant gap result?
  - High conductance/connectivity.
  - Metric induced rates.
- Additive gap for the complete graph is constant, i.e., $\textsc{cost}_{NA} - \textsc{cost}_{A} = \mathcal{O}(1)$.