Late points for random walks, and fluctuations of cover times

KU Math Stouffer Department Colloquium

Roberto Imbuzeiro Oliveira (IMPA)

Lawrence, Apr 25th 2012

[Joint work with Alan Prata from IMPA]

What is this talk about?

- What is random walk?

What is this talk about?

- What is random walk?
- What is the cover time?

What is this talk about?

- What is random walk?
- What is the cover time?
- What are late points?

What is this talk about? (II)

- Our contribution

What is this talk about? (II)

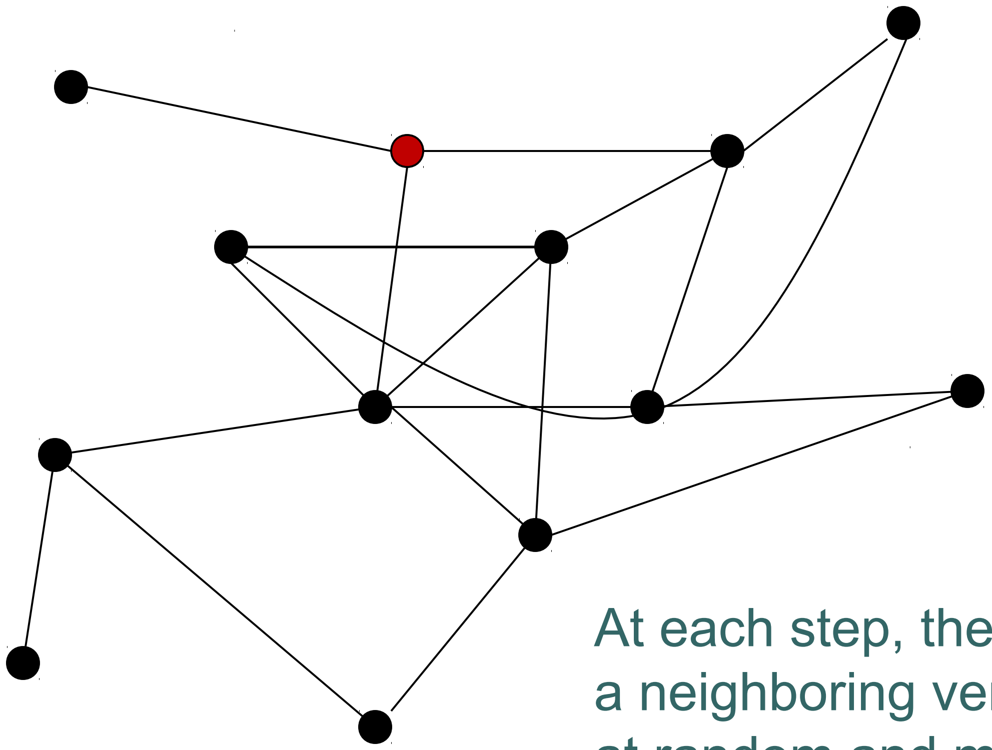
- Our contribution
- Main assuptions

What is this talk about? (II)

- Our contribution
- Main assuptions
- Key proof ideas, if there's time (blackboard)

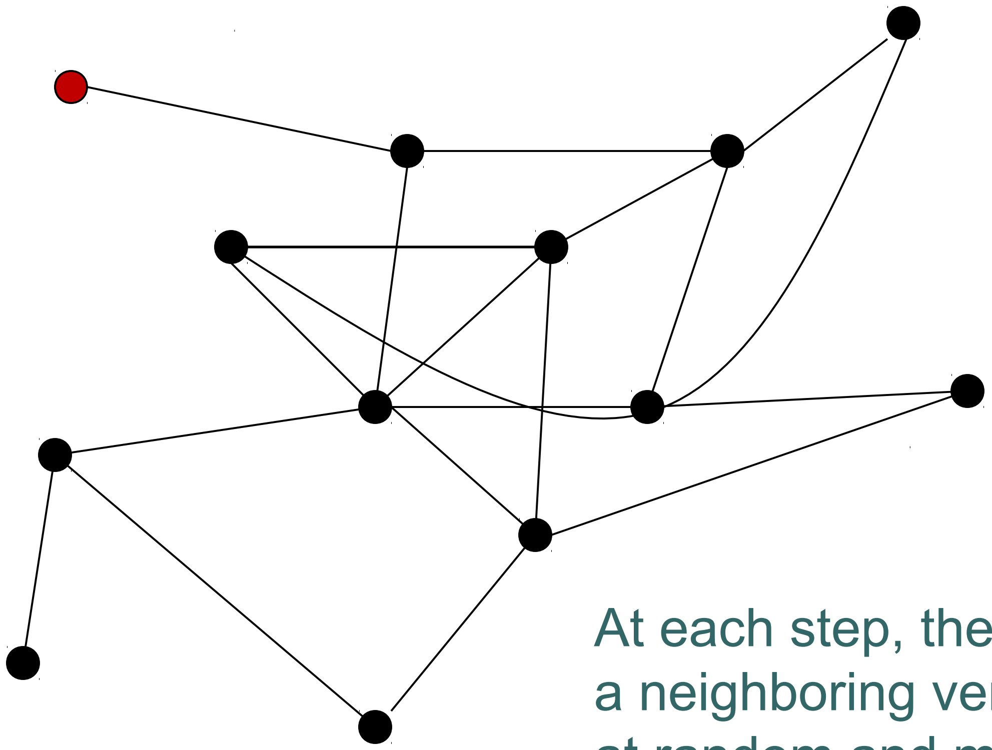
What is random walk over a finite graph?

RW on a finite graph



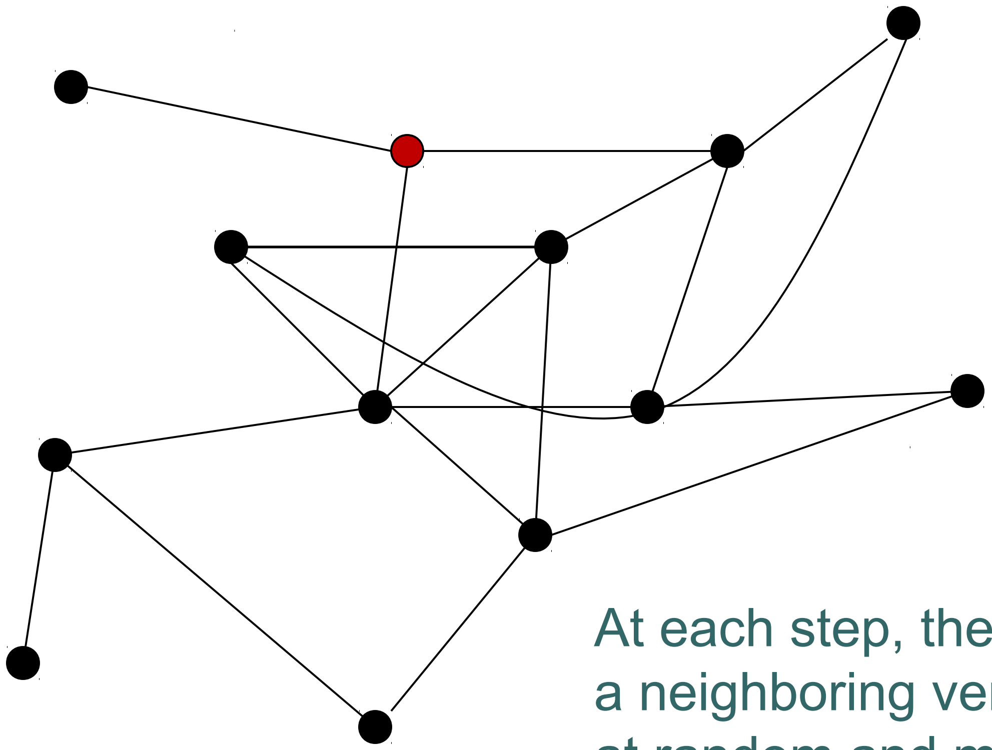
At each step, the red walker picks a neighboring vertex uniformly at random and moves there.

RW on a finite graph



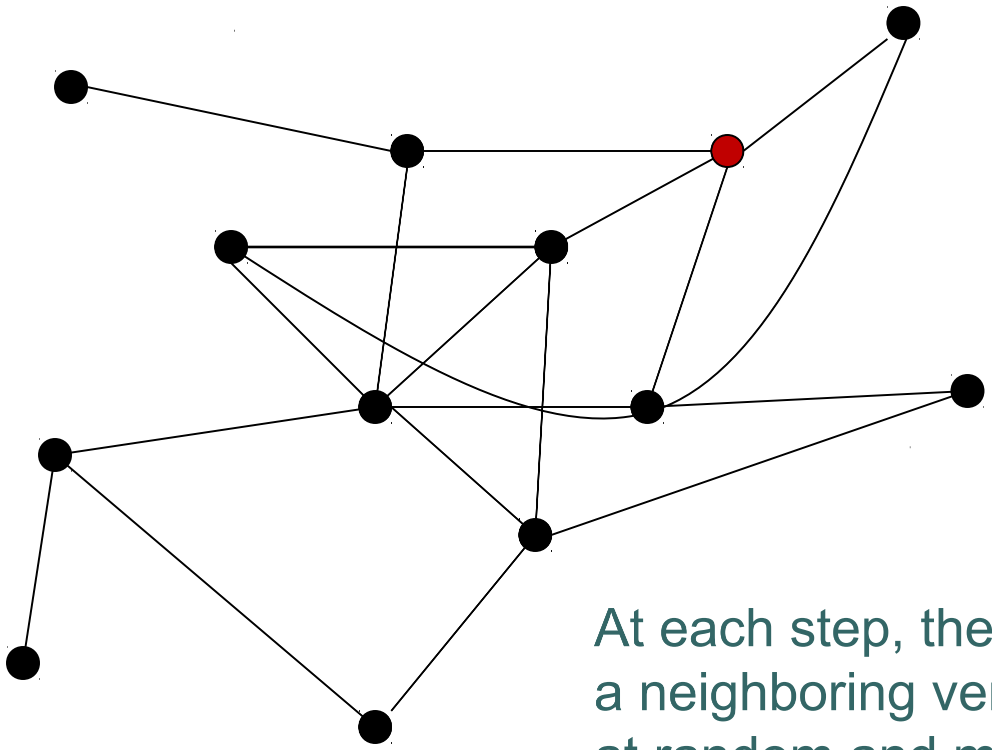
At each step, the red walker picks a neighboring vertex uniformly at random and moves there.

RW on a finite graph



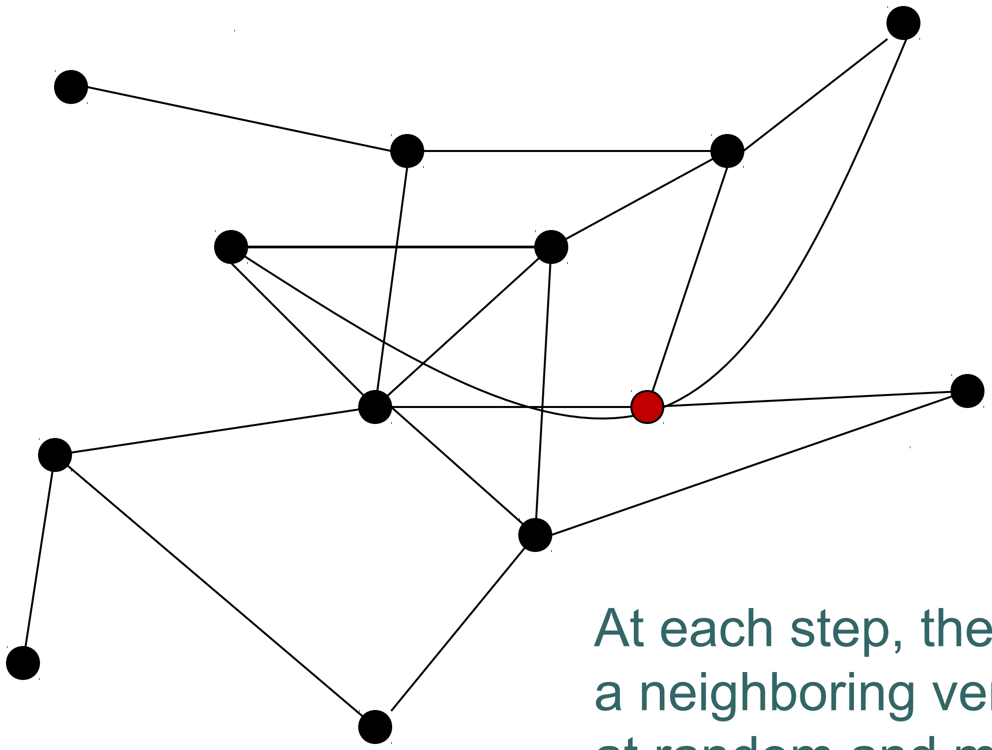
At each step, the red walker picks a neighboring vertex uniformly at random and moves there.

RW on a finite graph



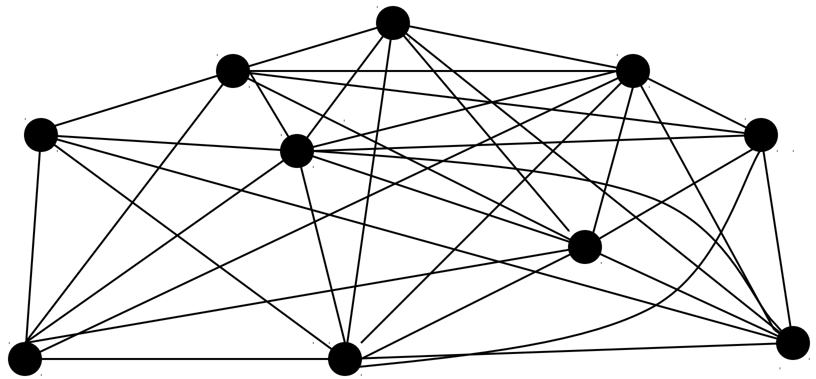
At each step, the red walker picks a neighboring vertex uniformly at random and moves there.

RW on a finite graph

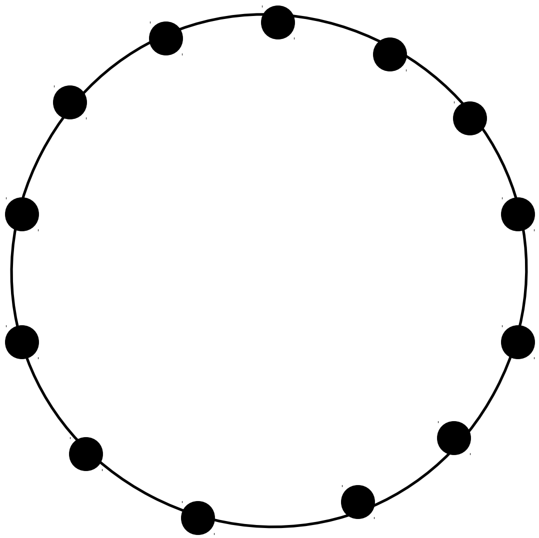


At each step, the red walker picks a neighboring vertex uniformly at random and moves there.

Examples

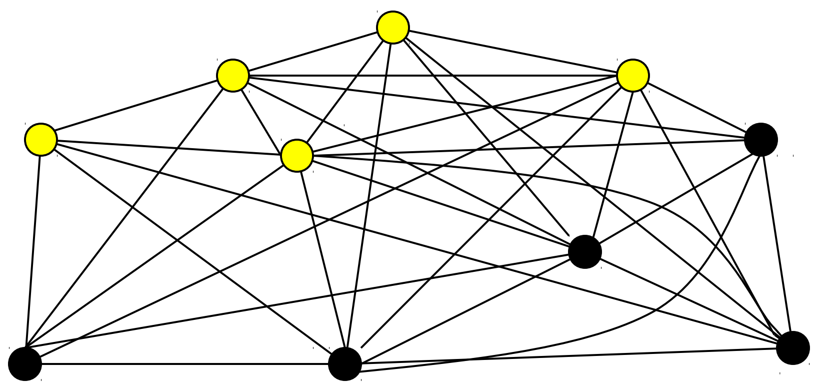


G = graph with n vertices and almost maximal number of edges

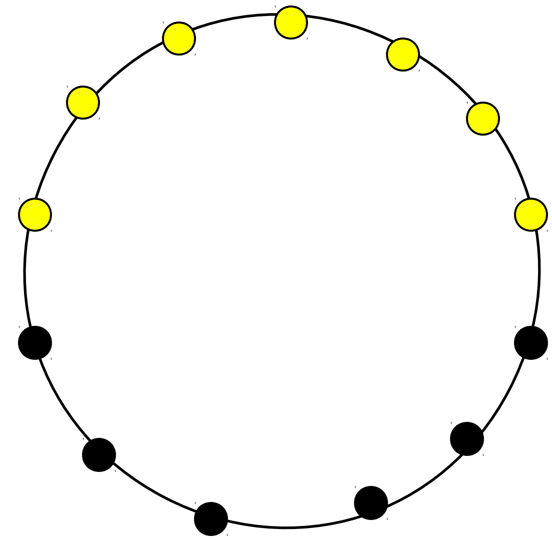


G = cycle with n vertices

RW can tell they're different

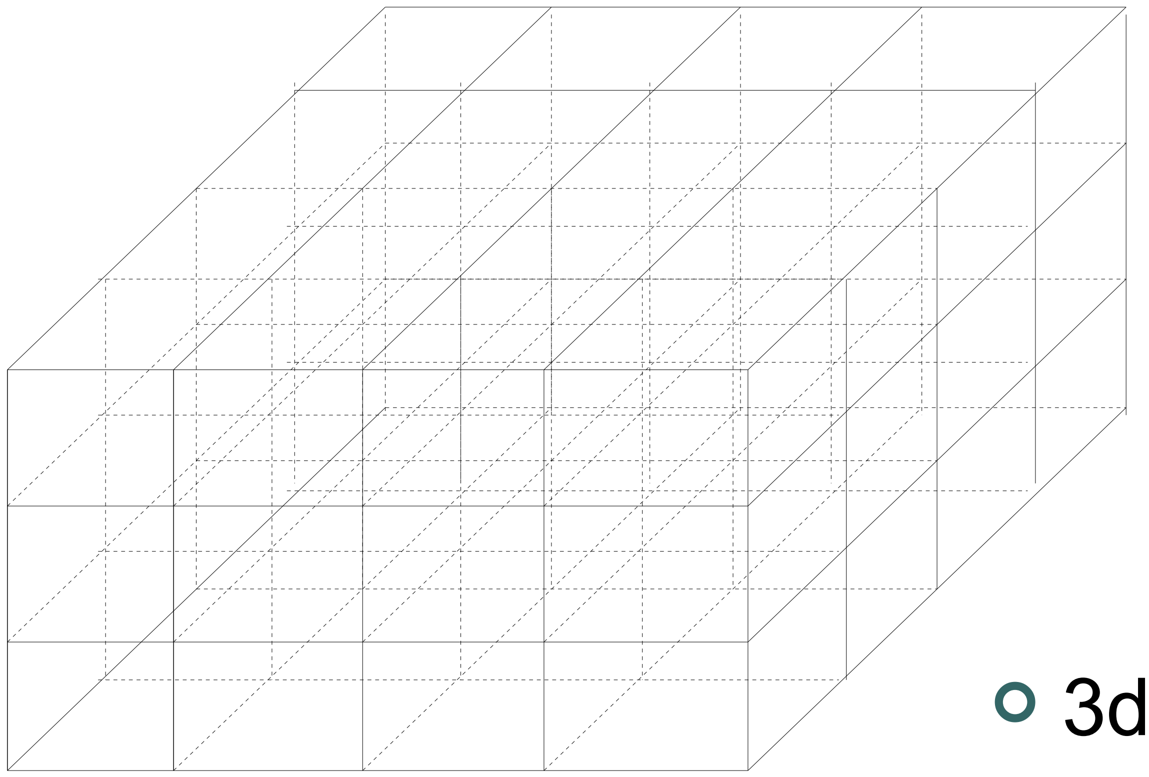


For the walker, yellow and black appear to be fair coin flips



Walker sees lots of yellow, then transition, then lots of black...

Another example



- 3d torus

The sliding window

HTTTTHHTHTHHTHTHTHH…

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH…

time 0

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH...

time 1

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH…

time 2

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH...

time 3

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH…

time 4

Repeated tosses of a fair coin

The sliding window

HTTTTTHHTHTHHTHTHTHH…

time 5

Repeated tosses of a fair coin

The sliding window

HTTTTTHHTHTHHTHTHTHH…

time 6

Repeated tosses of a fair coin

The sliding window

HTTTTHHTHTHHTHTHTHH…

time 7

Repeated tosses of a fair coin

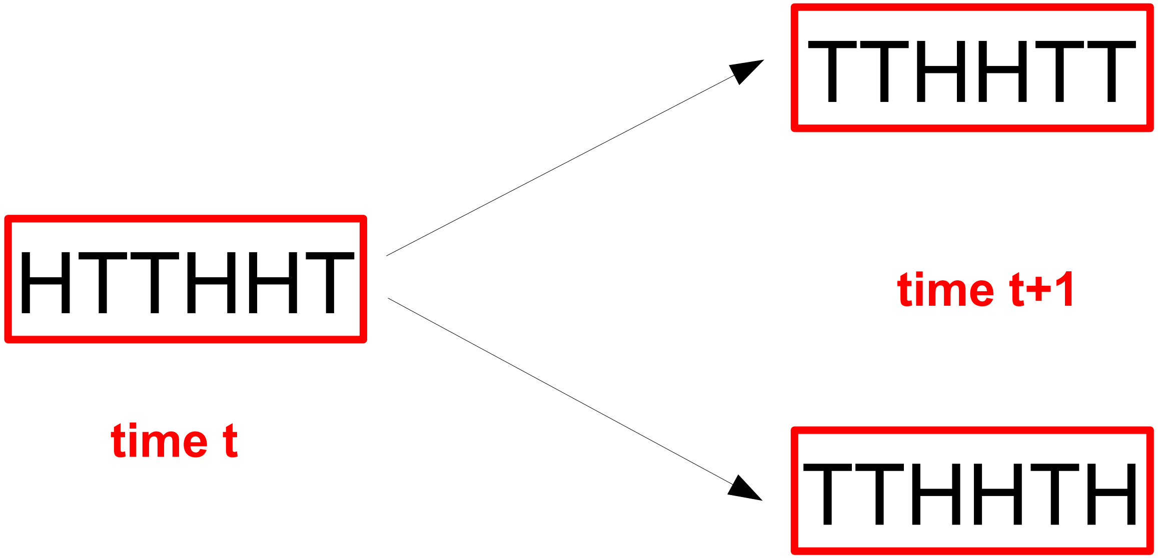
The sliding window

HTTTTHHT**HTHHTH**THTHH…

time 8

Repeated tosses of a fair coin

The sliding window

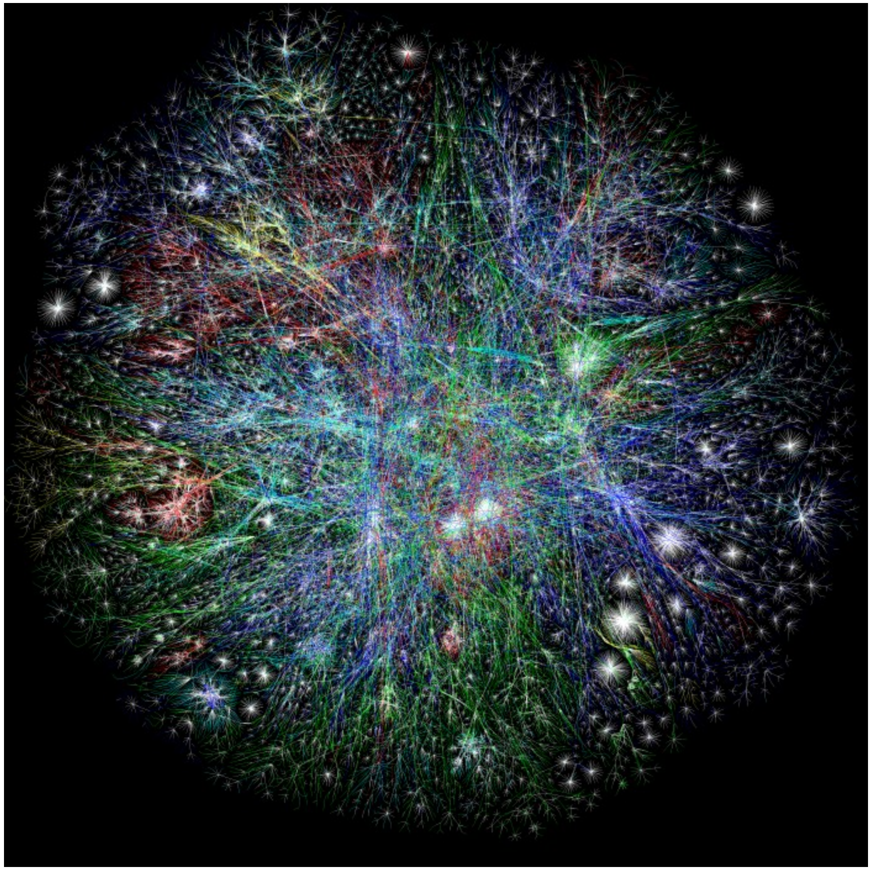


Random walk on an oriented graph

Variants of the basic model

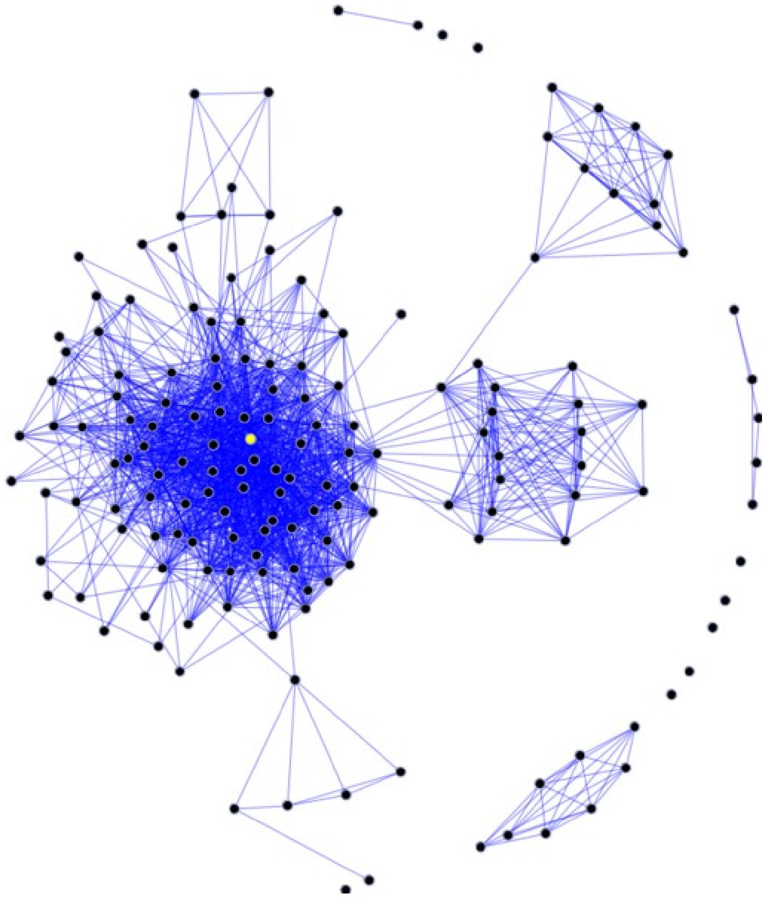
- Non-uniform transitions
- Directed transitions
- More general “Markov chains”

Why care?



Internet servers = vertices, connections = edges (Project OPTE)

A social network



People = vertices, “friendship” = edges

Google cares about RW

- The original Brin/Page paper on **Pagerank** features random walks and their ergodic properties quite proheminently.

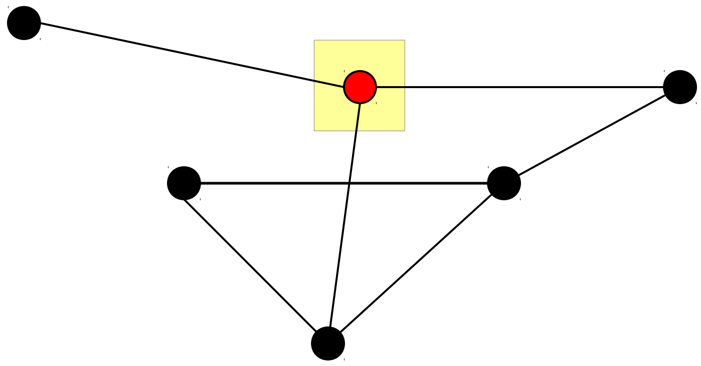
Other applications

- Many uses in
  Theoretical Computer Science

- Simulation of physical systems & Bayesian Statistics via
  Markov Chain Monte Carlo (MCMC)

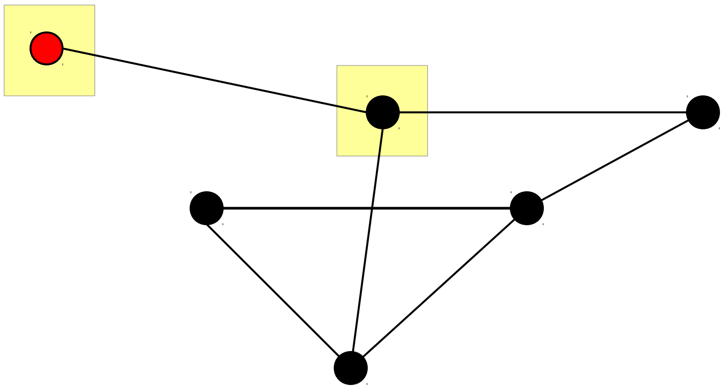
What are cover times ?

What is the cover time?



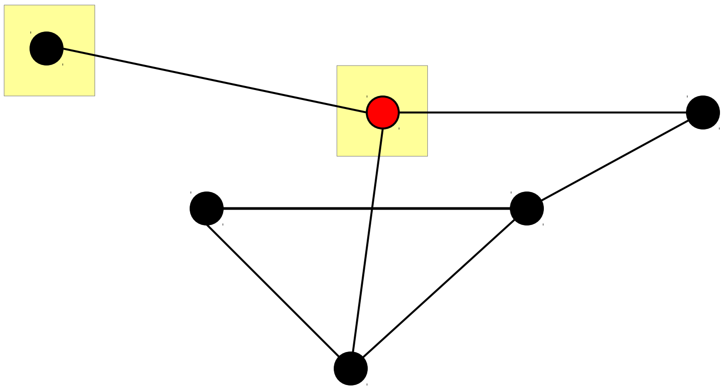
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



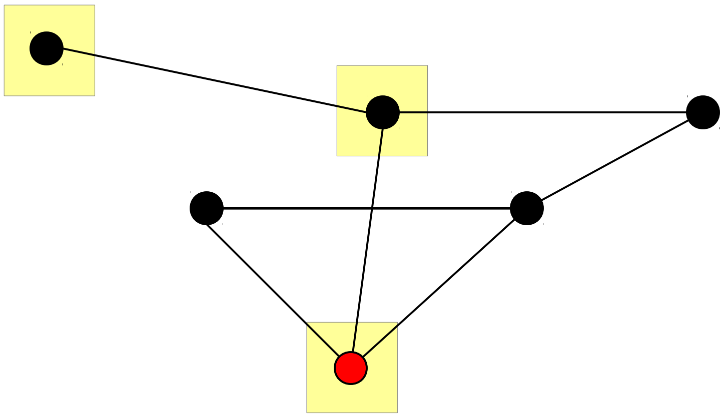
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



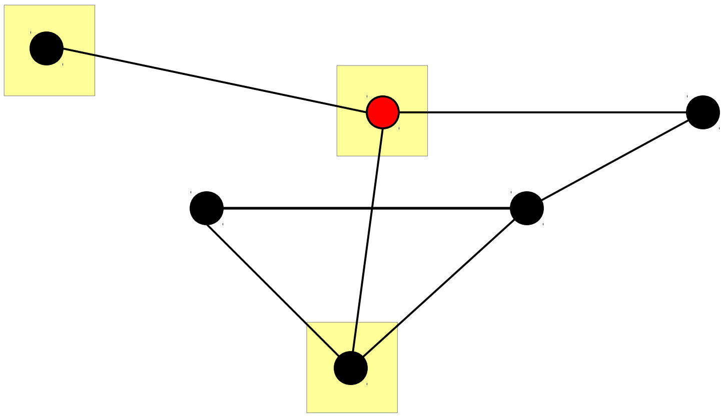
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



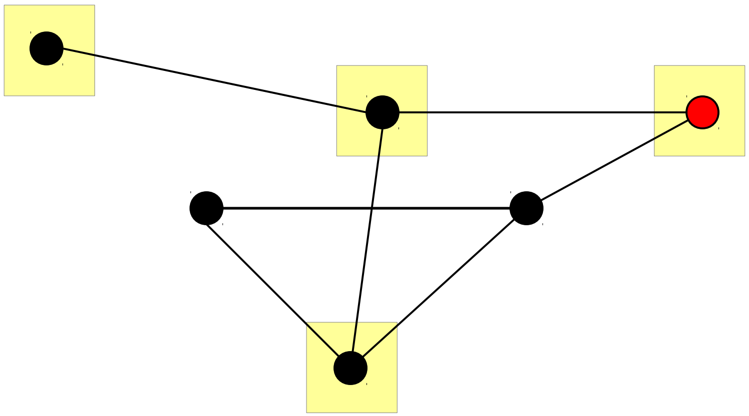
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



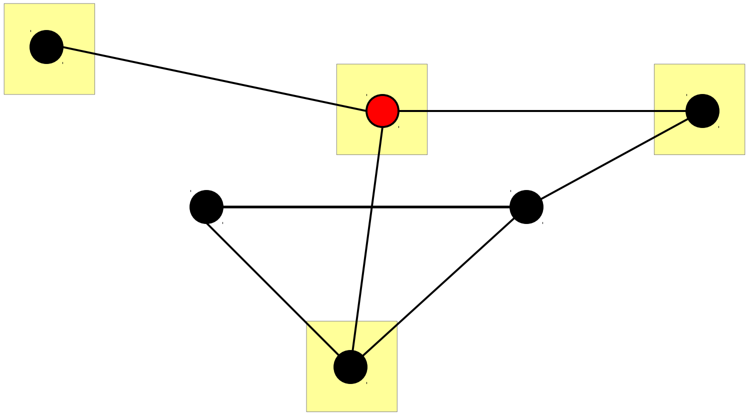
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



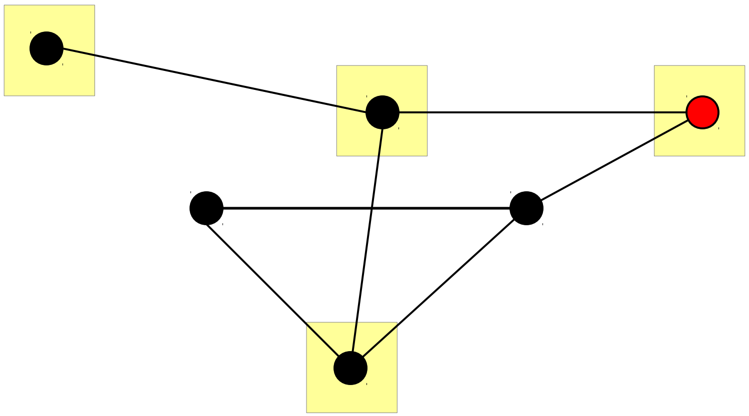
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



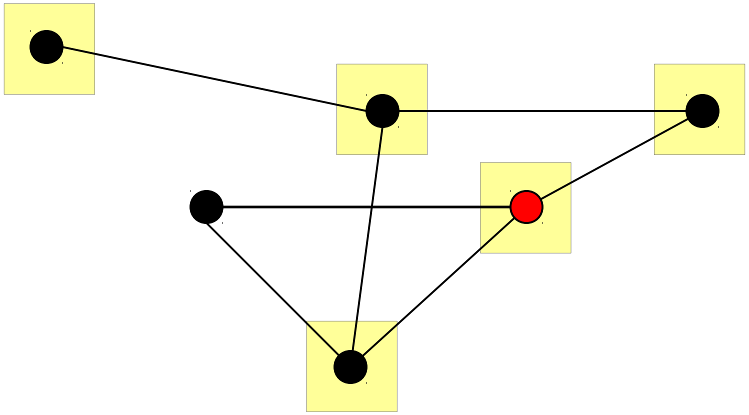
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



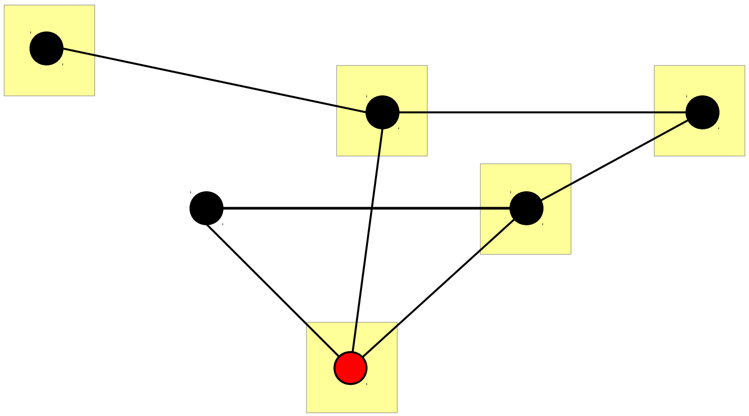
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



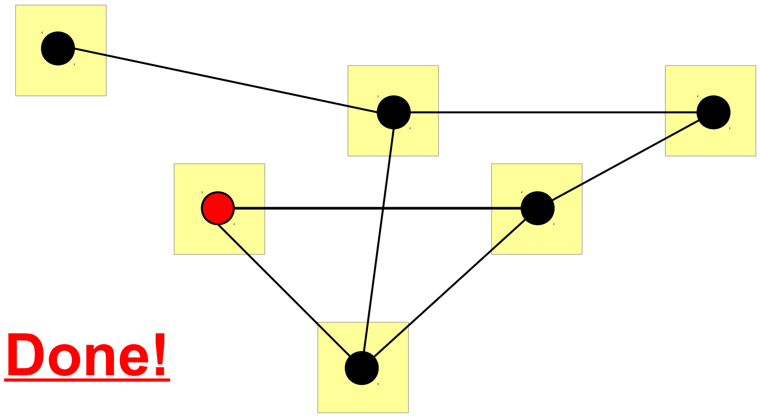
The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



The cover time is the time it takes for the red walker to visit every single vertex of the graph.

What is the cover time?



Done!

The cover time is the time it takes for the red walker to visit every single vertex of the graph.

Some notation

- **X(t)** = position of RW at time t
- **H(v)** = hitting time of vertex v,
  min {t : X(t) = v}
  (time of first visit to v)
- **C** = cover time = max{H(v) : v ϵ V}

Uses and properties

- **C** appears naturally when one thinks of RW as an exploration process, hence the CS interest.

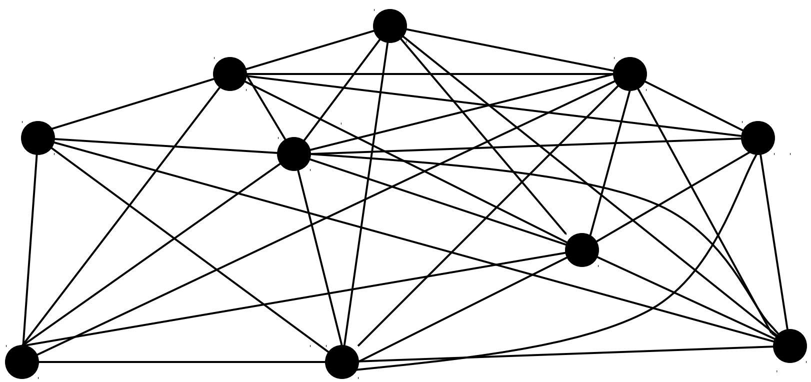
Uses and properties

- **C** appears naturally when one thinks of RW as an exploration process, hence the CS interest.
- **C** for sliding windows is essentially the time to see all patterns of size w.

Expectations

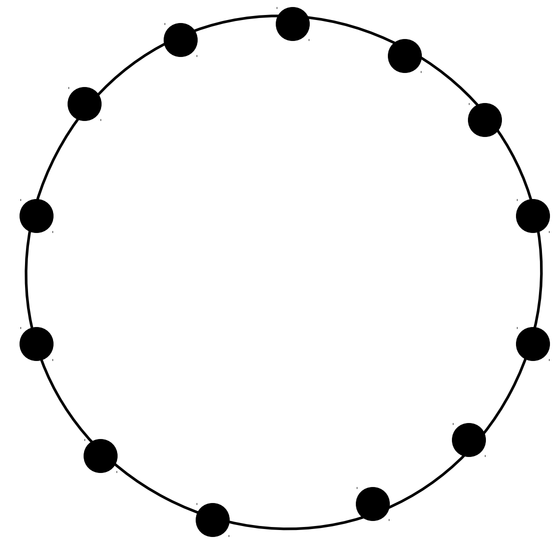
- A lot is known about the **expectation** of **C**, in general and in concrete examples.

Two examples



G = graph with n vertices and almost maximal number of edges

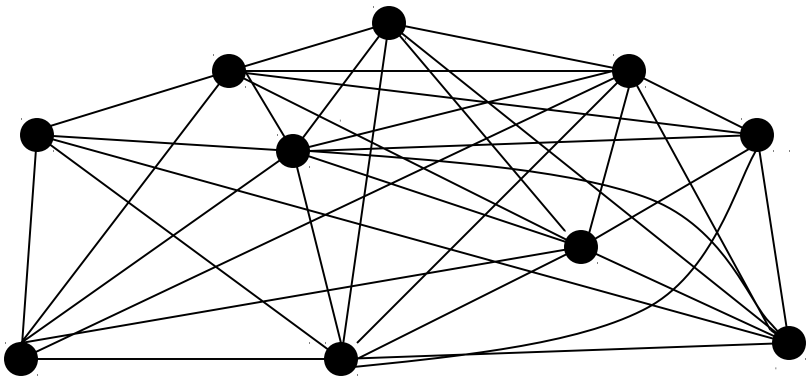
Ex(C) = n log n + (error)



G = cycle with n vertices

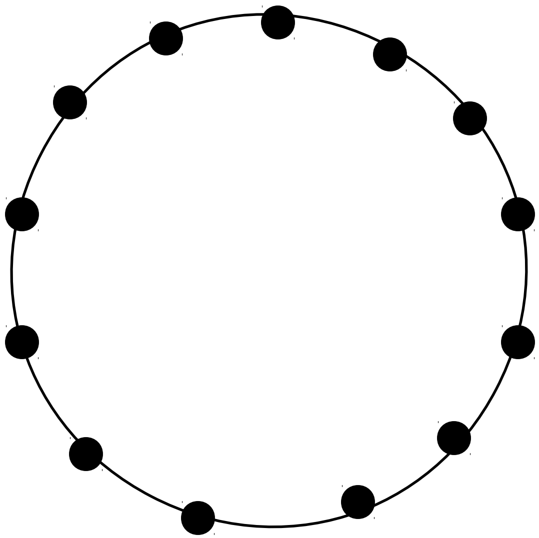
Ex(C) = n^2

Two examples



G = graph with n vertices and almost maximal number of edges

Ex(C) = n log n + (error)



G = cycle with n vertices

Ex(C) = n^2

Extremes

- What are the largest/smallest possible values of **Ex(C)** over all unoriented graphs with **n** vertices and all possible starting points?

Extremes

- What are the largest/smallest possible values of **Ex(C)** over all unoriented graphs with **n** vertices and all possible starting points?

  Answer: $(4/27)\ n^3$ and $n \log n$

  [Brightwell & Winkler/Feige]

The Gaussian Free Field

- Ding/Lee/Peres, Ann. Math. (to appear)

Ex(C) is “proportional” to ...

$|E| \, \mathrm{Ex}(\max\{\mathbf{G(v)} : v \in V\})^2$

The Gaussian Free Field

Gaussian vector with G(X(0))=0 &
covariances = effective resistances

Ex(C) is “propo al” to ...

$|E|\ \mathrm{Ex}(\max\{\mathbf{G(v)} : v \in V\})^2$

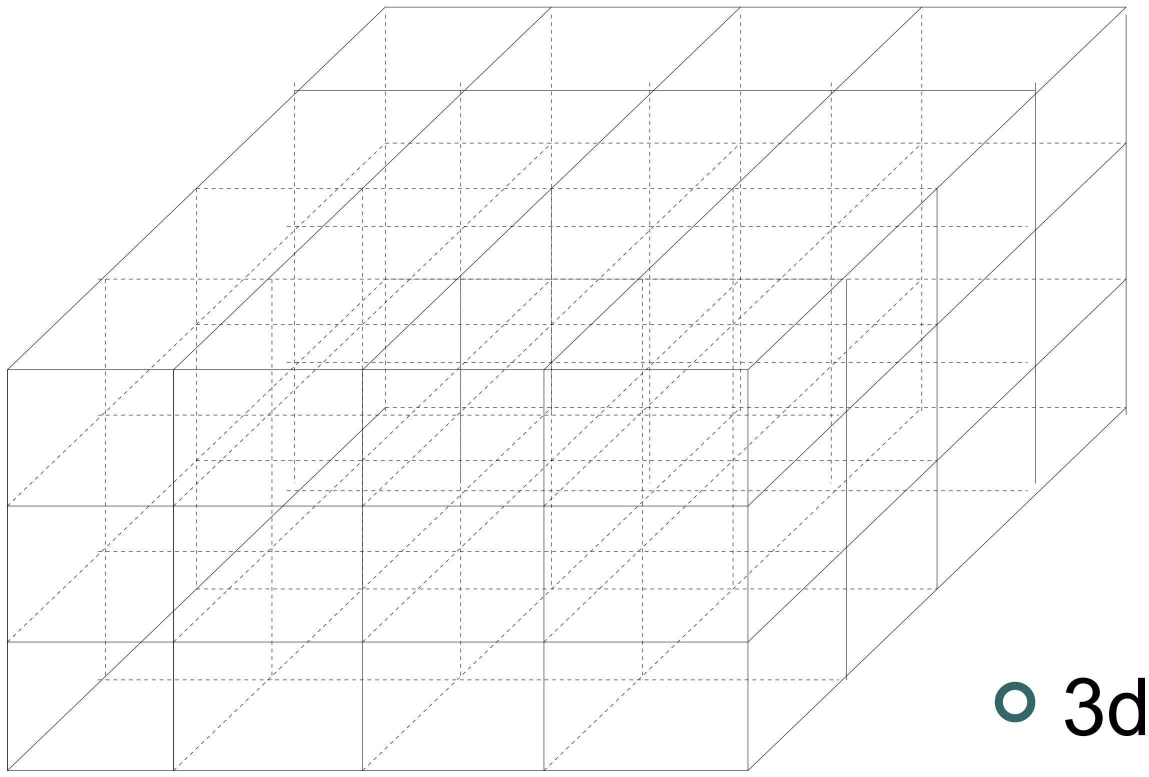
Weak laws of large numbers

- For many large graphs it is known that **C** is tighly concentrated around

$$\mathbf{Ex(C) \sim g(G)\ n \log n}$$

where g(G) is a constant (general concentration result by Aldous'90).

Weak laws of large numbers



- 3d torus

  ~ g(d) n log n

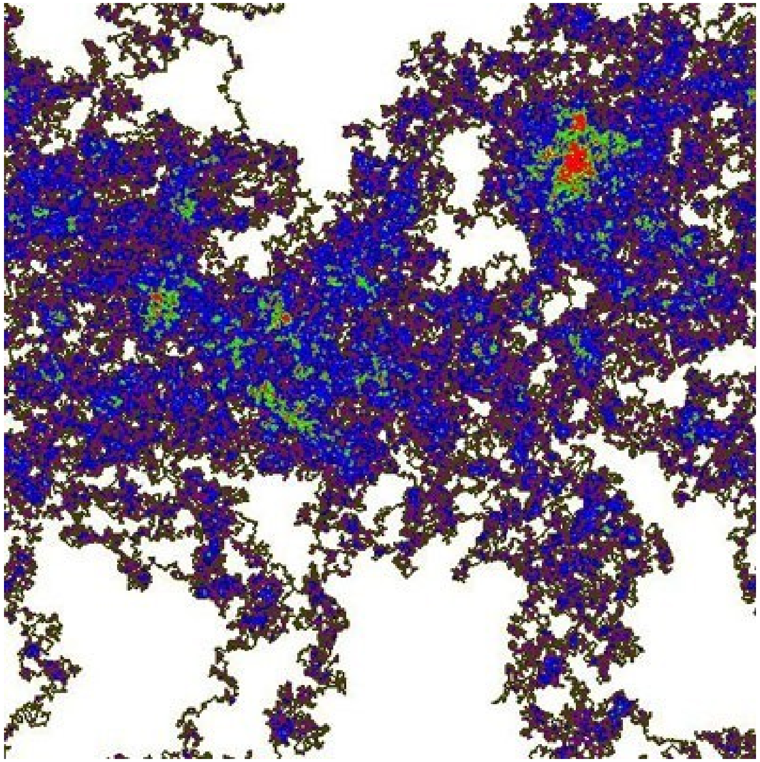
Weak laws of large numbers

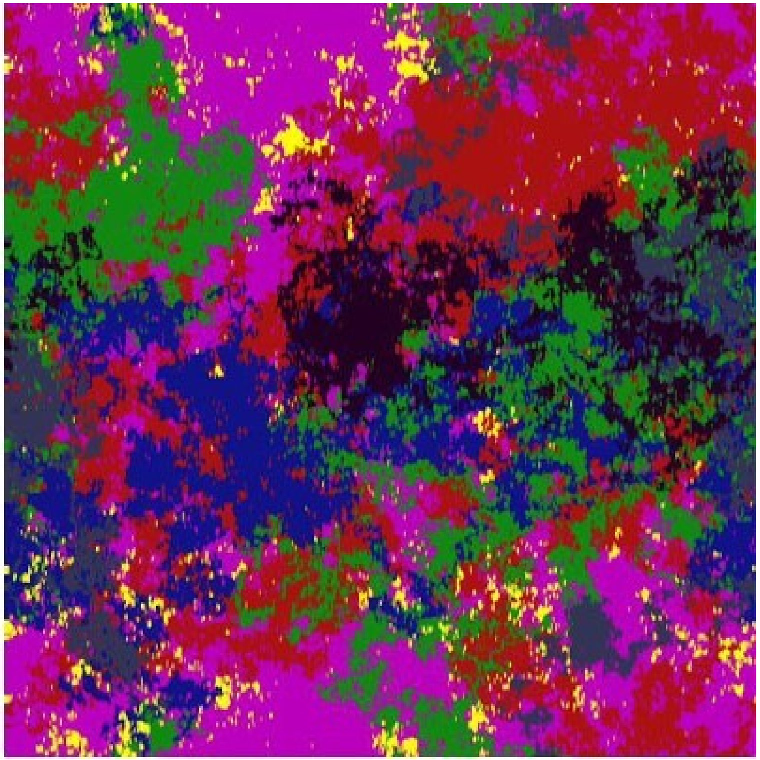
- Other examples of

  $C \sim g(G)\ n \log n$

  include hypercubes ,sliding windows typical regular graphs, expander graphs, ...

What about 2d?





Dembo, Peres, Rozen and Zeitouni: 2001/4

Fluctuations of cover times

What are fluctuations?

- **Law of Large Numbers:** $C/a(n) \sim 1$
- **Fluctuations:** behavior in distribution of

$$\frac{C - a(n)}{s(n)}$$

(think "Central limit theorem")

A folklore conjecture

- **Gumbel law fluctuations:**

  For "high dimensional graphs":

  $\Pr[C - g(G)\, n \ln n < \mathbf{c}\, g(G)\, n] \to \exp(-e^{-\mathbf{c}})$

A folklore conjecture

- **Gumbel law fluctuations:**

  For "high dimensional graphs":

  $\Pr[C - g(G)\, n \ln n < \mathbf{c}\, g(G)\, n] \to \exp(-e^{-\mathbf{c}})$

  (For the complete graph, $g(G)=1$.)

Known instances

- **Gumbel law fluctuations:**

  Hypercubes [Devroye & Sbihi'90]

Known instances

- **Gumbel law fluctuations:**

  Hypercubes [Devroye & Sbihi'90]

  Sliding windows over coin tosses [Móri'90]

Known instances

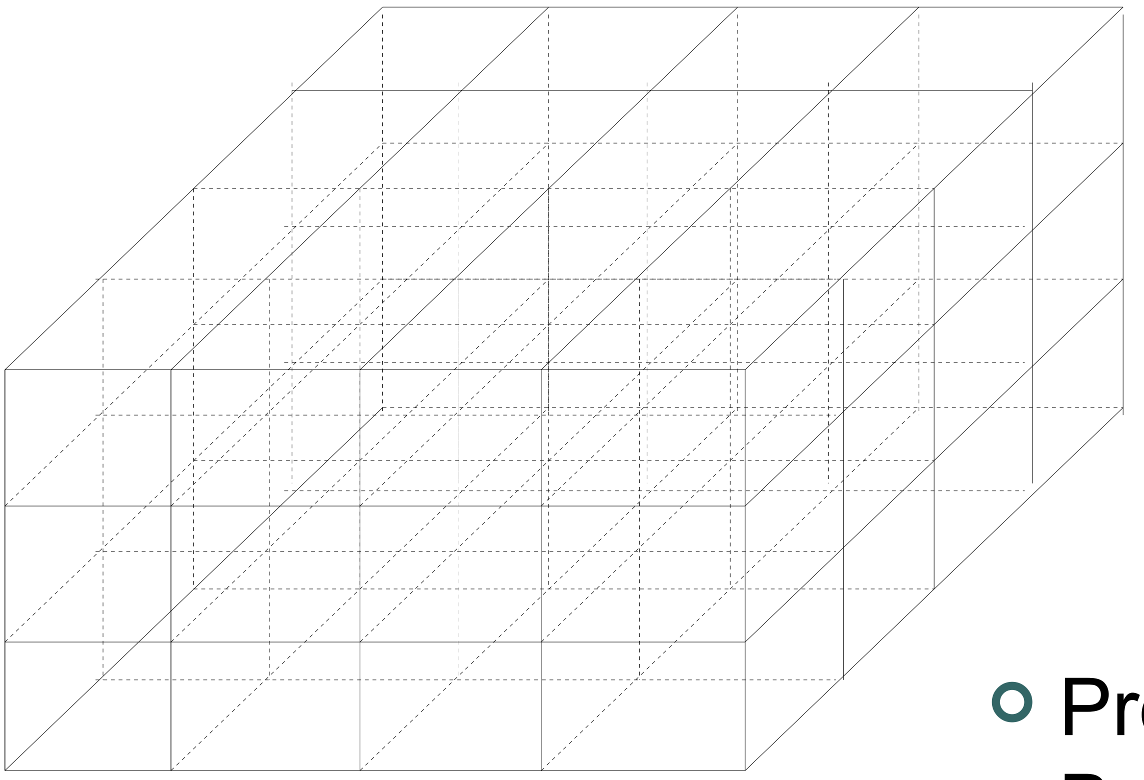
- **Gumbel law fluctuations:**

  Hypercubes [Devroye & Sbihi'90]

  Sliding windows over coin tosses [Móri'90]

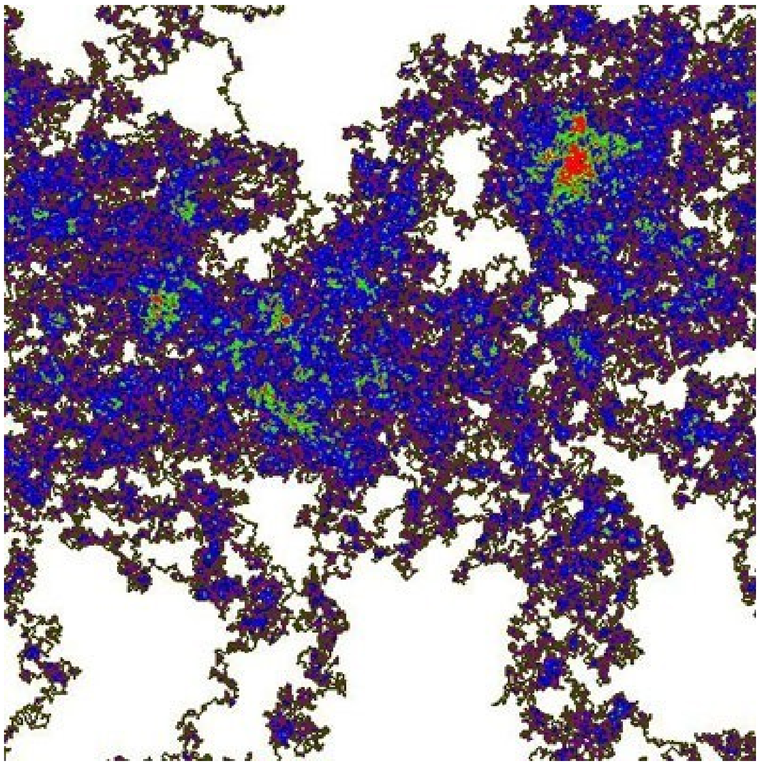
  Graphs where all vertices are “far” [Aldous'94]

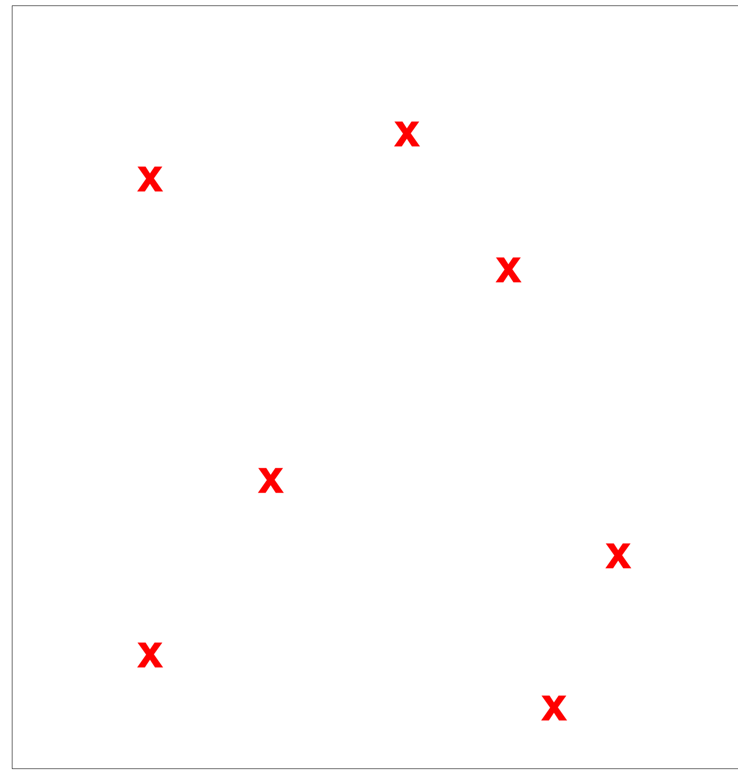
The 3d torus case



- Preprint by Belius (2012)

Two vs three dimensions





Poisson limit in the 3d case!

Poisson limit in 3d

- **Approx. law of uncovered set at time**

$$T(\mathbf{c}) = g(d)\, n \log n + \mathbf{c}\, g(d)\, n$$

Po(e^{-c}) points thrown uniformly at random in the d-dim. cube

Poisson limit in 3d

- **Approx. law of uncovered set at time**

$$T(\mathbf{c}) = g(d)\, n \log n + \mathbf{c}\, g(d)\, n$$

Po(e^{-c}) points thrown uniformly at random in the d-dim. cube

(Gumbel law follows from this!)

Late points, and our contribution

Our contribution

Joint work with Alan Prata (PhD, IMPA)

Our contribution

Joint work with Alan Prata (PhD, IMPA)

- The Gumbel law is **general**

Our contribution

Joint work with Alan Prata (PhD, IMPA)

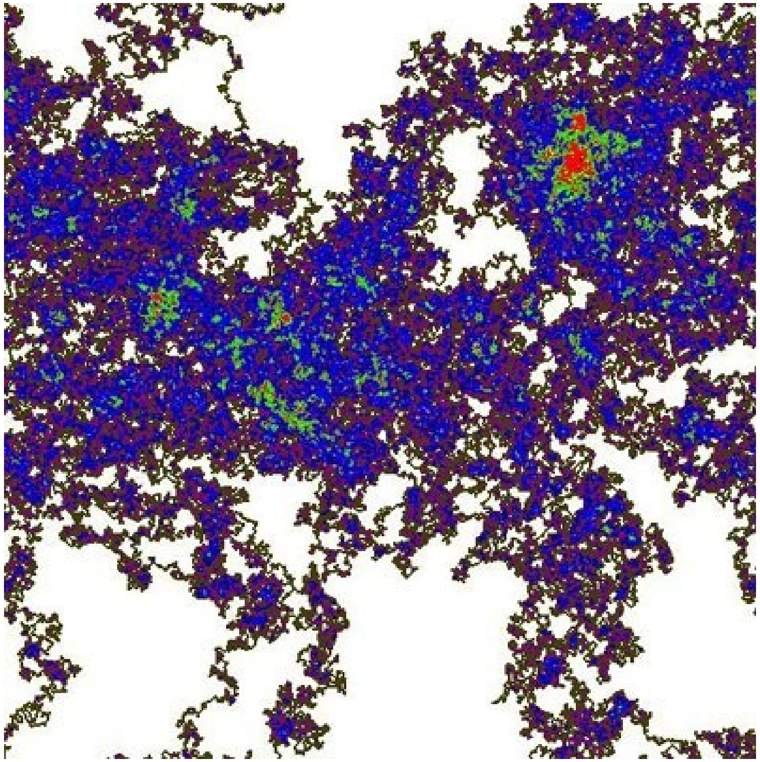
- The Gumbel law is **general**
- The Gumbel law is a consequence of properties of **late points**

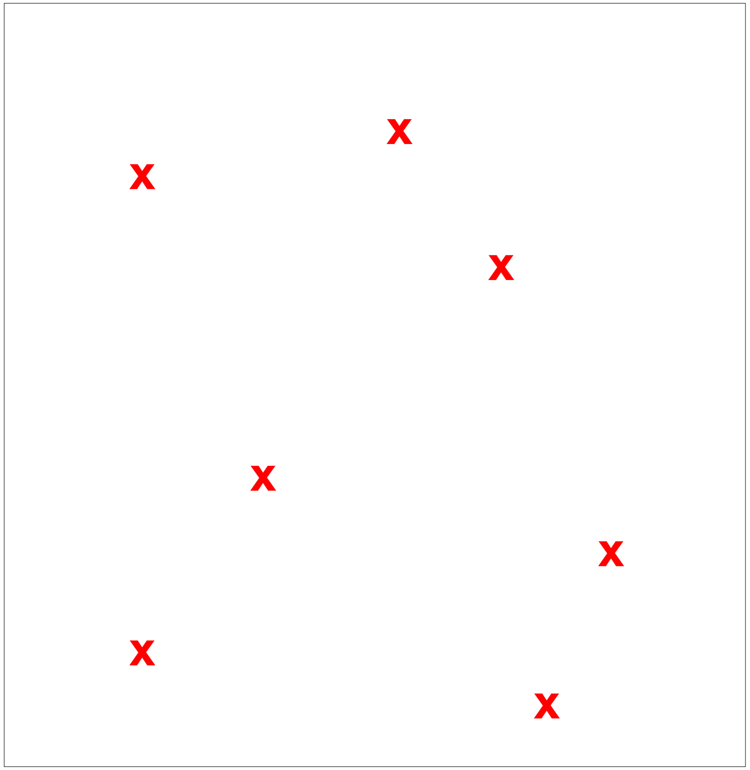
Our contribution

Joint work with Alan Prata (PhD, IMPA)

- The Gumbel law is **general**
- The Gumbel law is a consequence of properties of **late points**
- General thm. about the **evolution of** the set of **uncovered points**

It's all about late points





General Poisson limit theorem

A sense of the main thm

- Recall the folklore conjecture:

$$\Pr[C < g(G)\, n \log n + c\, g(G)\, n] \approx \exp(e^{-c})$$

A sense of the main thm

- Recall the folklore conjecture:

$$\Pr[\mathbf{C} < g(G)\, n \log n + \mathbf{c}\, g(G)\, n] \approx \exp(e^{-\mathbf{c}})$$

- We prove a stronger theorem about the uncovered set U(**c**) at this time.

A sense of the main thm

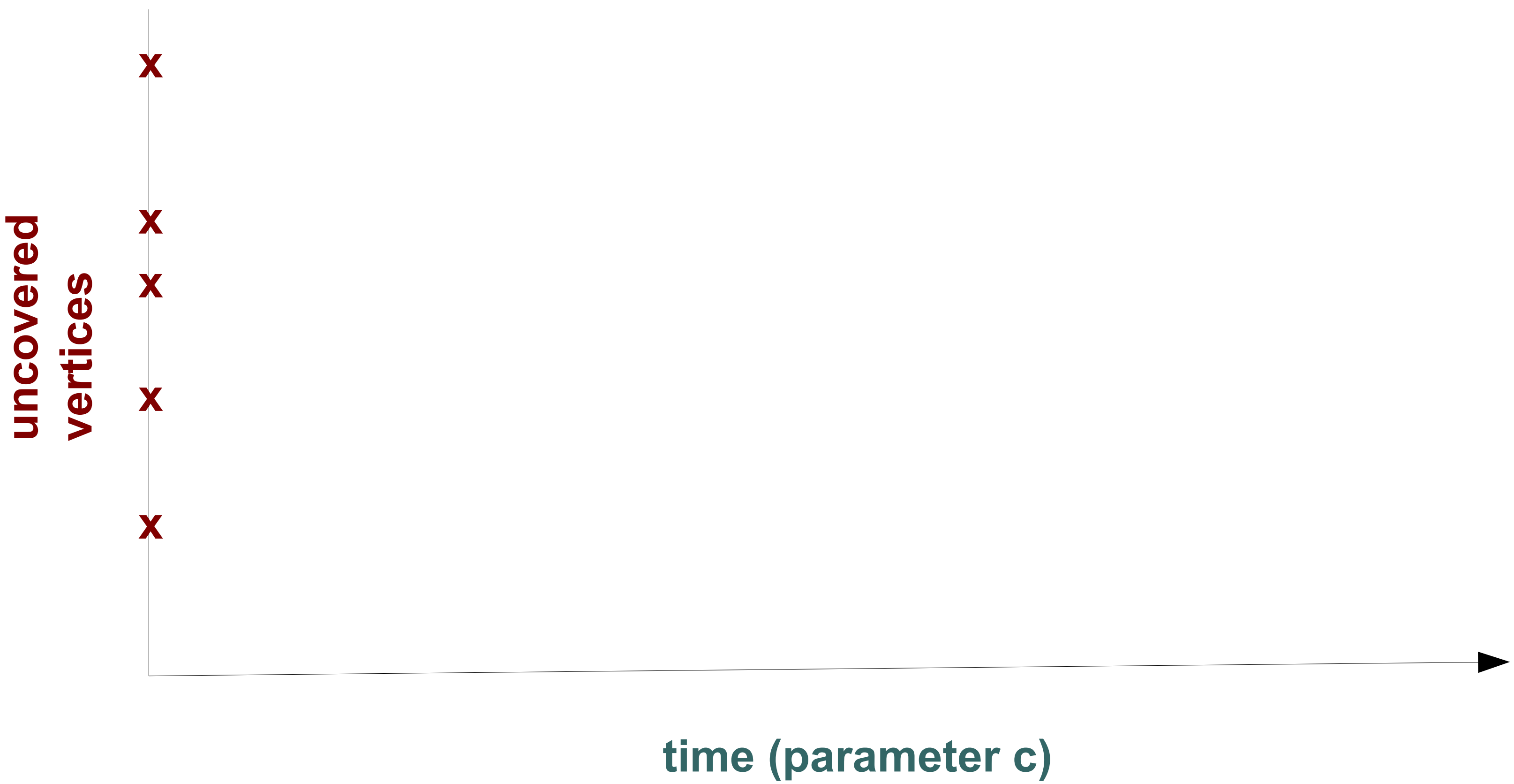
- Recall the folklore conjecture:

$$\Pr[\mathbf{C} < g(G)\, n \log n + \mathbf{c}\, g(G)\, n] \approx \exp(e^{-\mathbf{c}})$$

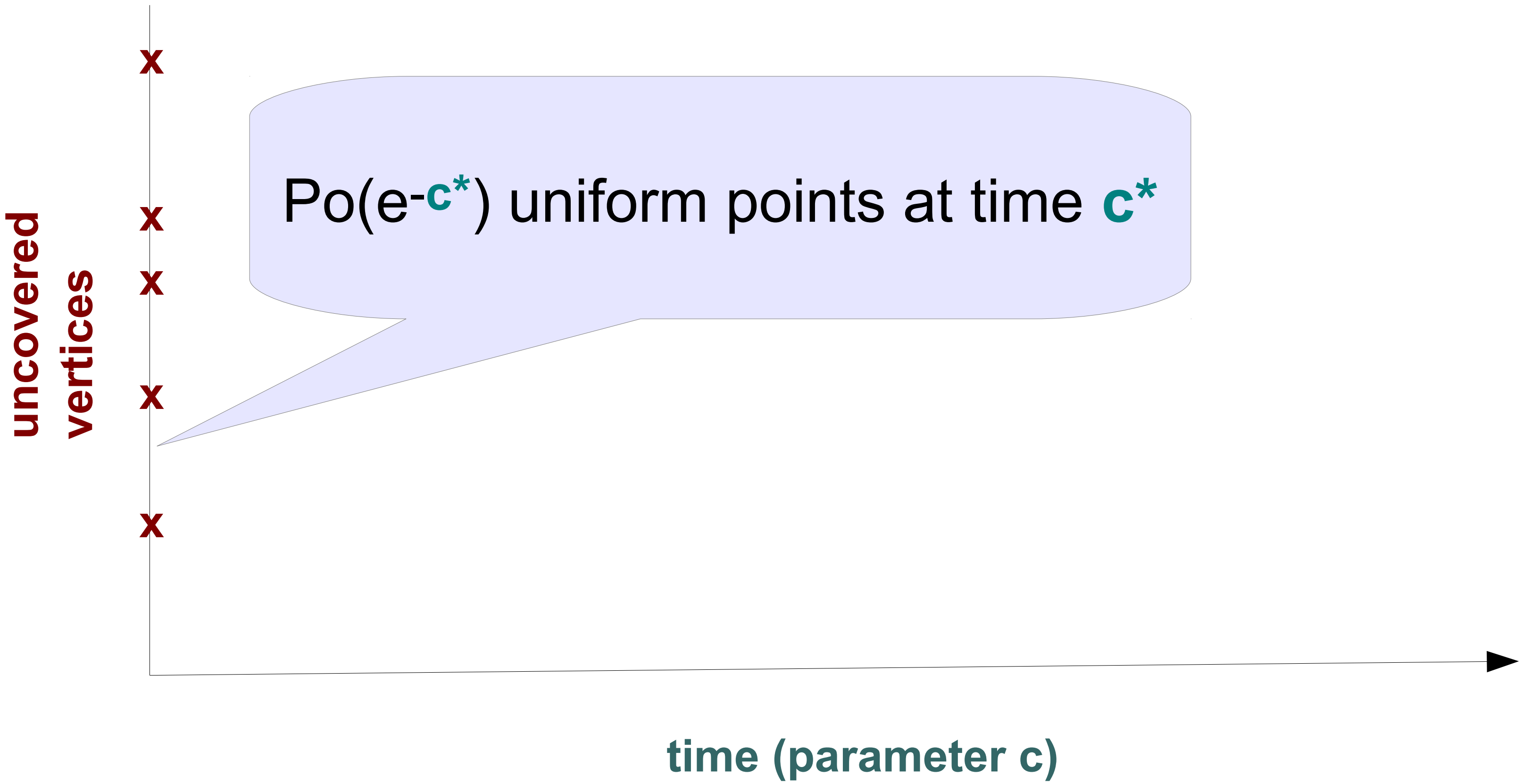
- We prove a stronger theorem about the uncovered set U(**c**) at this time.

U(c) ~ Po($e^{-\mathbf{c}}$) points, uniform over V

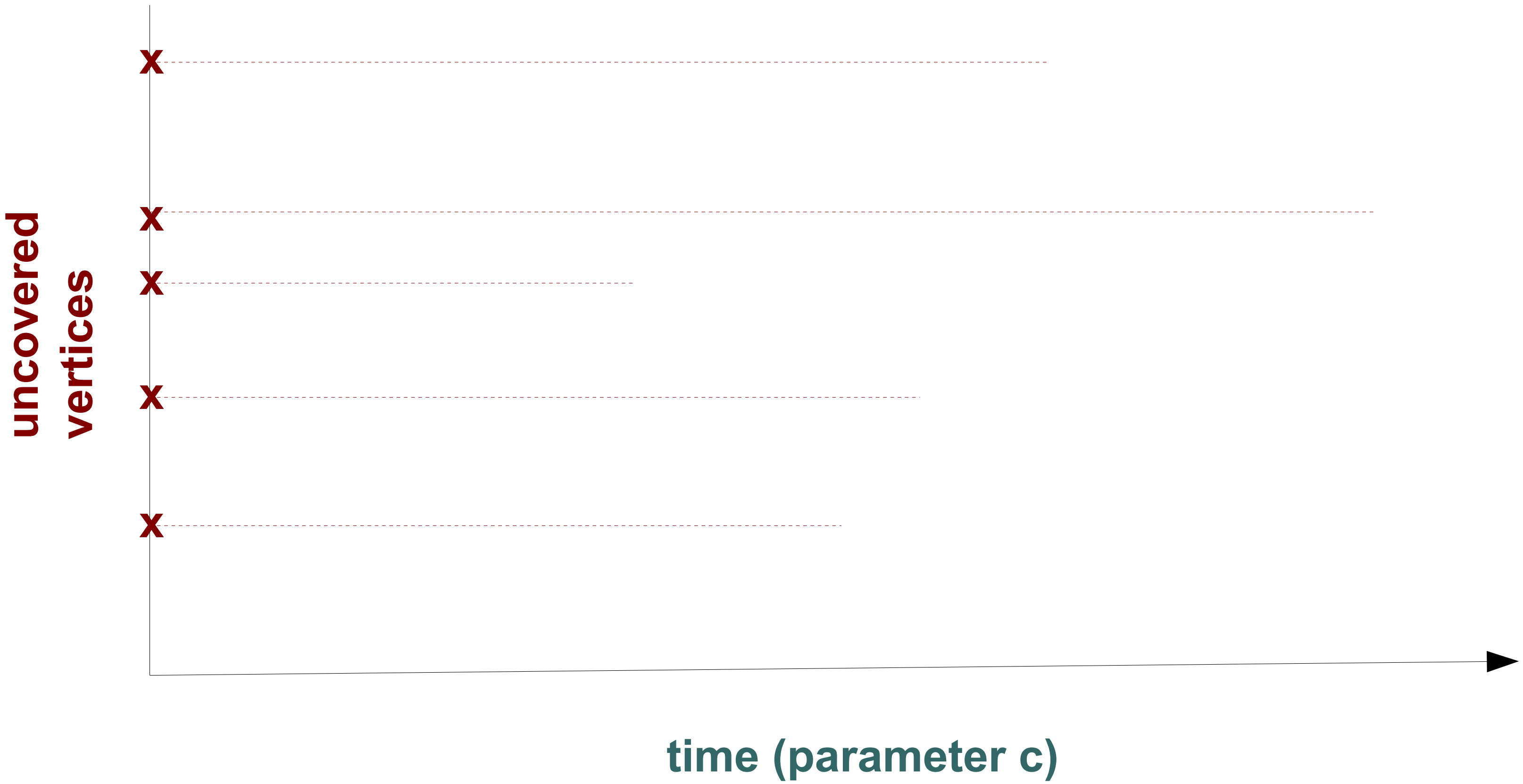
A sense of the main thm (II)



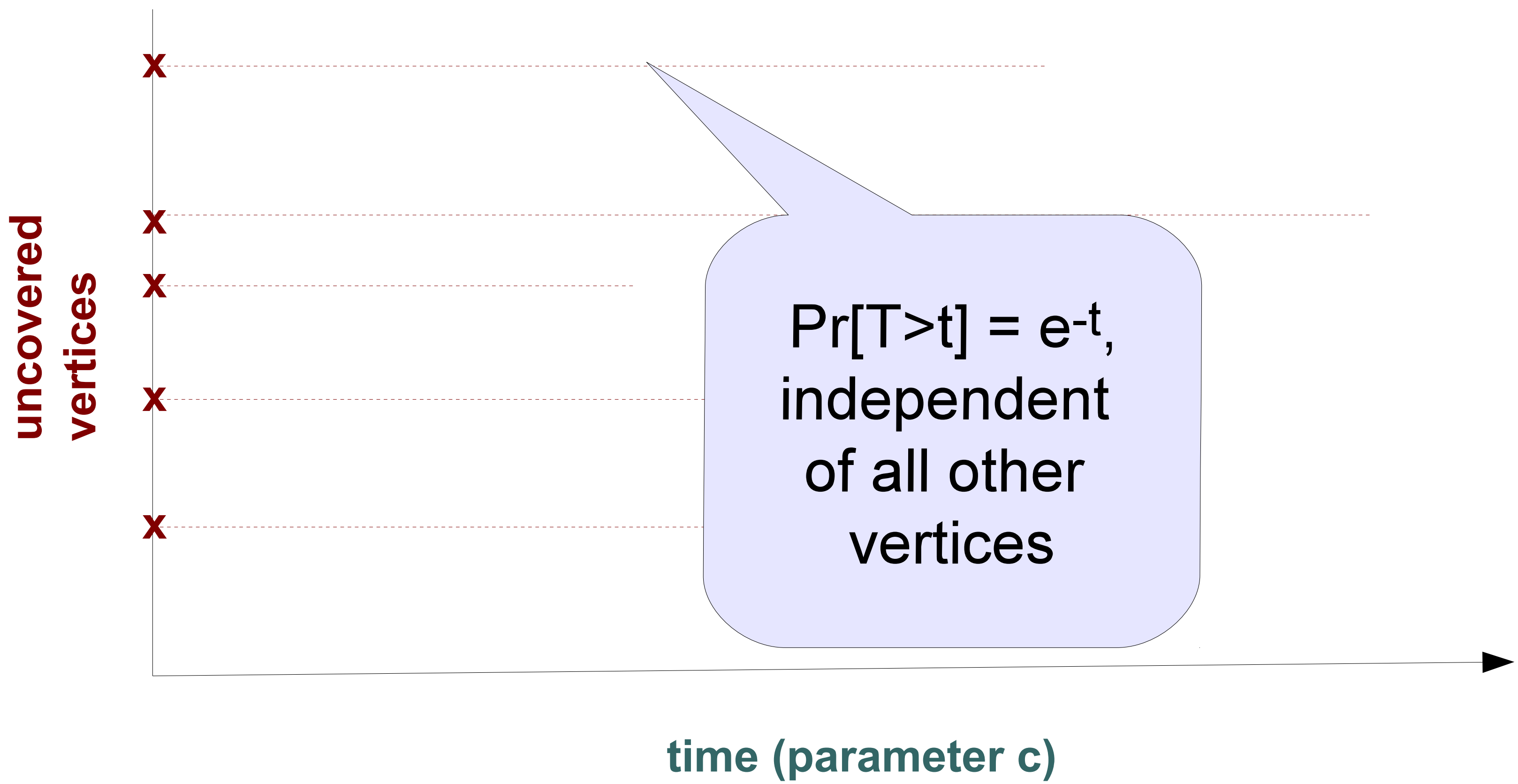
A sense of the main thm (II)



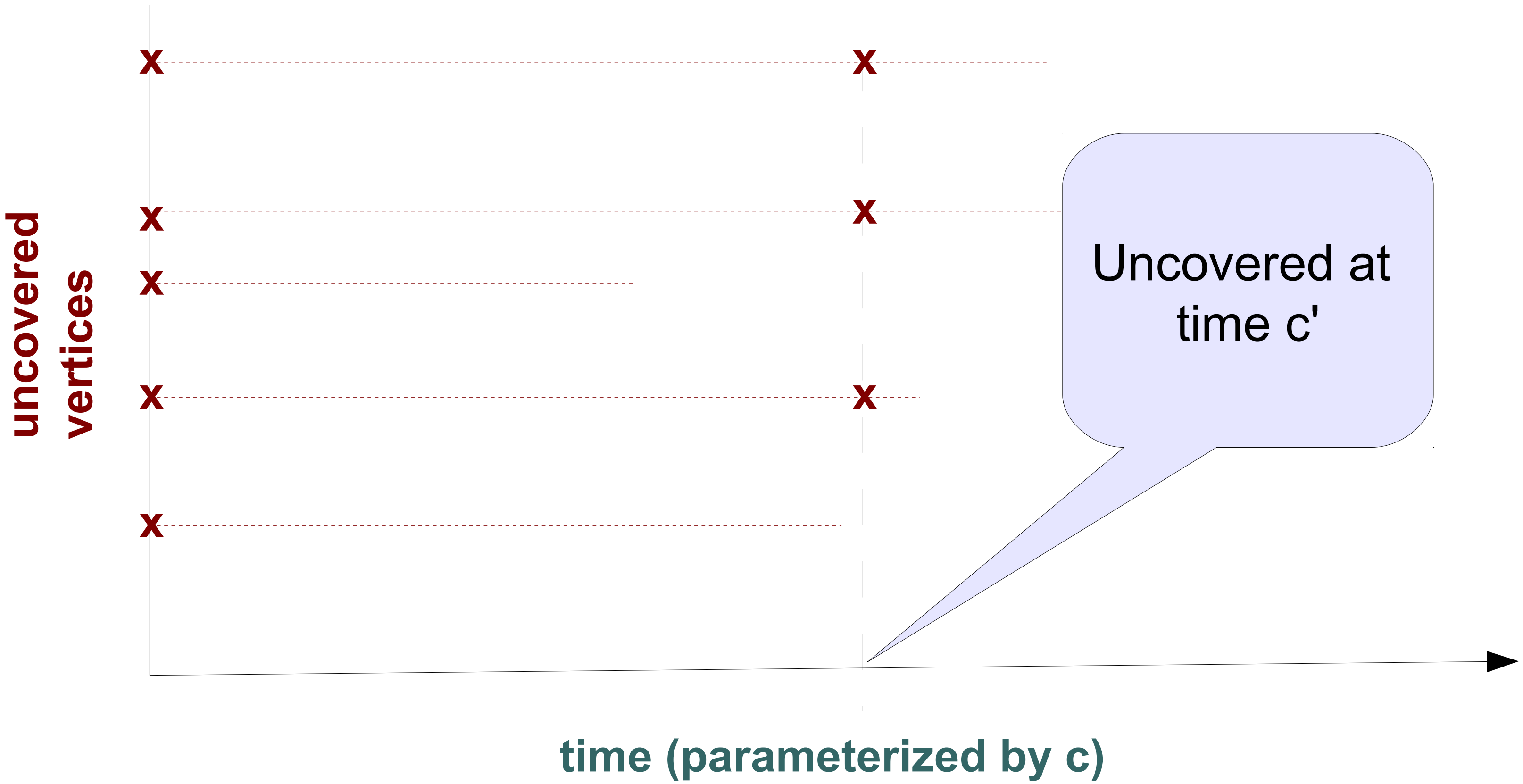
A sense of the main thm (II)



A sense of the main thm (II)



A sense of the main thm (II)



An important corollary

- For any **k**, the last **k** points to be covered are nearly uniformly distributed over the subsets of size **k** of the vertex set.

  (Conjectured for k=1 by Aldous in 1992)

Main result in a nutshell

- $Po(e^{-c})$ uniform uncovered points at fixed time

$$g(G)\ n \log n + c\ g(G)\ n$$

- Later evolution: indenpendent exponentials
- Last k points are essentially uniform

The assumptions: fast mixing + local transience

Two technicality

- We need to work with random walks whose stationary distributions are uniform, or nearly so.
- We sometimes need to work with lazy walks (at each step stay put w/ prob. 1/2)

Assumptions (I)

- **Def:** **T** is a mixing time for RW on G if

  for all starting points $X(0) = x$ and all sets of vertices A,

$$\Pr[X(T) \in A] < |A|/n + |A|/n^2$$

- Need $T < n^a$ where $a<1$ (fast mixing).

Assumptions (II)

- Recall that the cover time is the maximum of hitting times:

  $H(v) = \inf \{ t : X(t)=v\}$

- We need that nearly all vertices have the same hitting time **g(G) n**, up to a small error, when starting from uniform position.

Assumptions (III)

- Let **H(v,w)** denote the minimum of H(v) and H(w)

  H(v,w) = min {t : X(t)=v or X(t)=w}

- Need that **Ex(H(v,w)) < (1-r) g(G) n,r>0** (hitting 2 points strictly faster than 1)

Assumptions in a nutshell

- Expected hitting times are nearly all nearly equal
- Hitting time of 2 strictly smaller than hitting time of 1
- Most vertices are far from each other
- Fast mixing (sometimes requires lazyness)

Assumptions (IV)

- Say vertices v and w are "close" if:

$$\Pr[H(w)<T|X(0)=v] > 1/(\log n)^2$$

or vice versa.

- Need that for all v at most $n^{o(1)}$ vertices are close.

Now on to the blackboard

References

Main reference

- These results are in the second part of Alan Prata's thesis, soon to appear in preprint form. (Should be soon available soon from IMPA's website.)

References: useful books

- Aldous/Fill "Reversible Markov Chains and Random Walks on Graphs" (book "in preparation" since 1990!)
- Levin, Peres e Wilmer "Markov Chains and Mixing Times" (AMS; also online)
- U. Feige : slides on cover times in CS

References

- J. Ding/J. Lee/ Y. Peres, Ann. Math (to appear – Gaussian free field)
- D. Aldous, J Theo Prob 3(4), 1990 (weak law of large numbers for C/Ex(C))
- Dembo/Peres/Rosen/Zeitouni, Ann. Math., 160 (2004) 433—464 (2d torus)

References: Gumbel law

- Devroye & Sbihi, J Theo Prob vol 3 (4), 1990
- Aldous and Fill (see above), chapters on cover times and transitive graphs
- D. Belius, arXiv: 1202.0190 (3d torus)