Exact distributions of cover times for N independent random walkers in one dimension

Sanjib Sabhapandit
Raman Research Institute, Bangalore

collaborators: Satya N. Majumdar & Gregory Schehr
Université Paris-Sud, Orsay, France

What is cover time ?

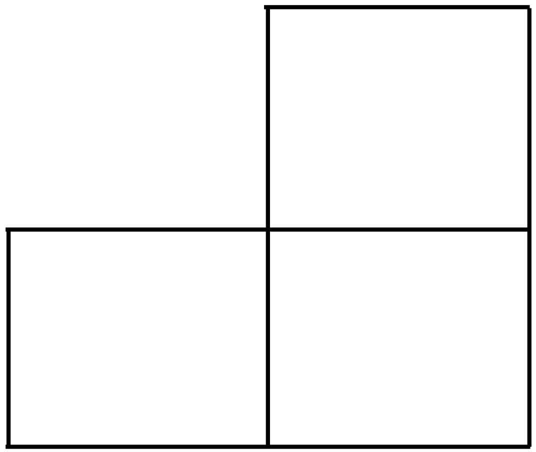
(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

What is cover time ?

(for a random walk)

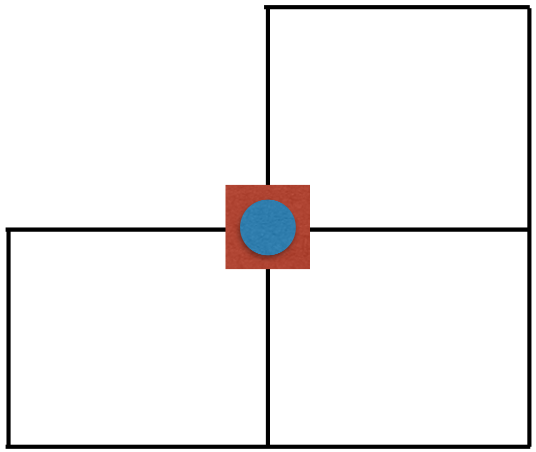
It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.



What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

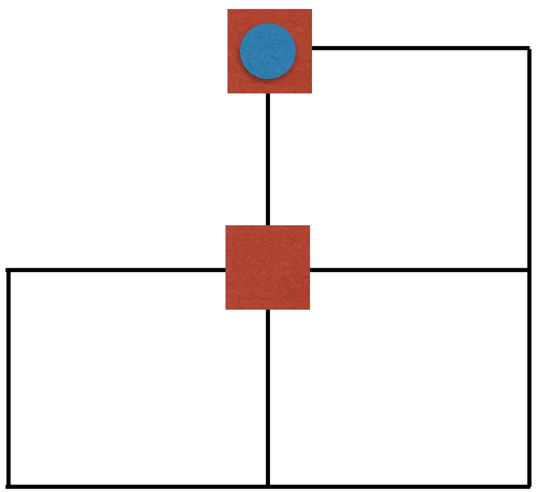


Time = 0

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

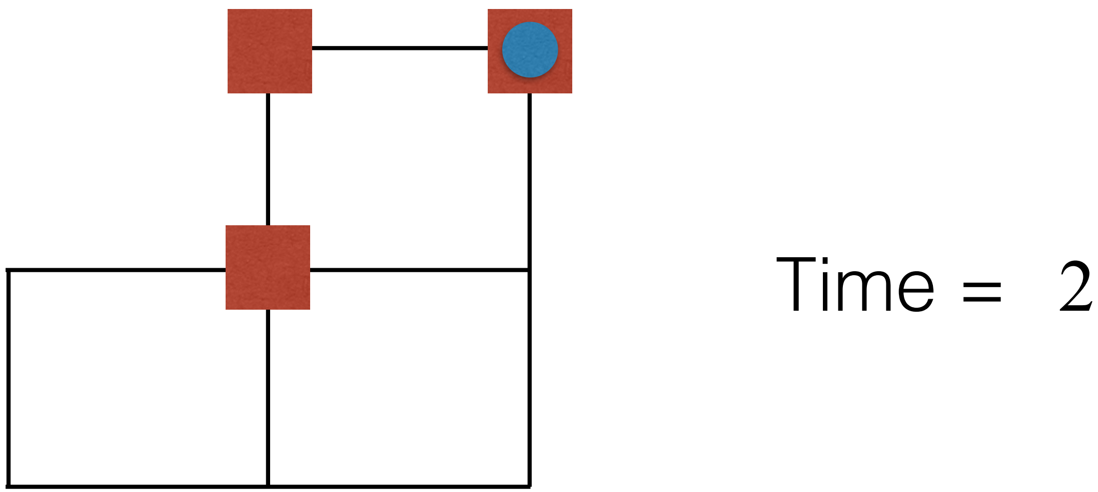


Time = 1

What is cover time ?

(for a random walk)

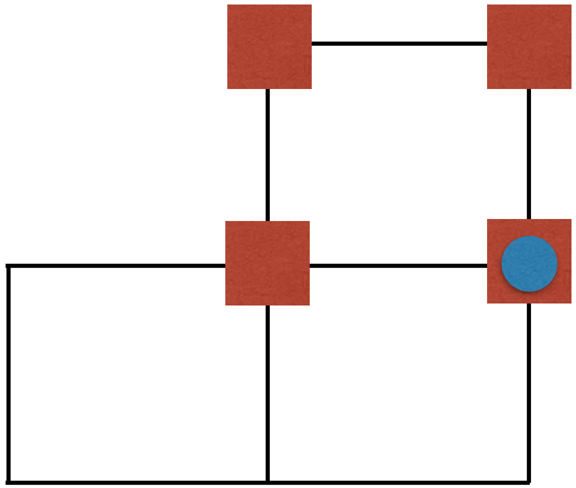
It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.



What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

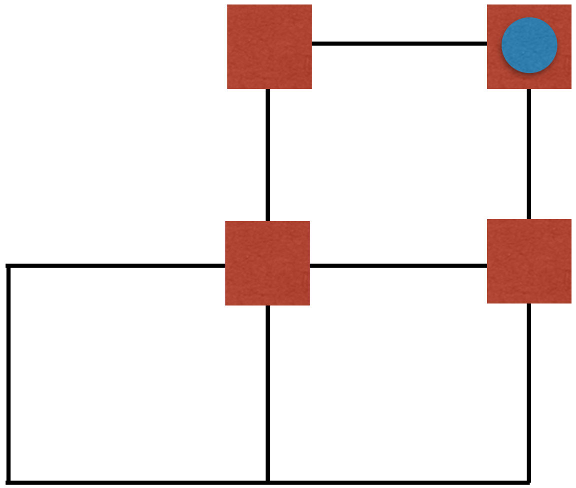


Time = 3

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

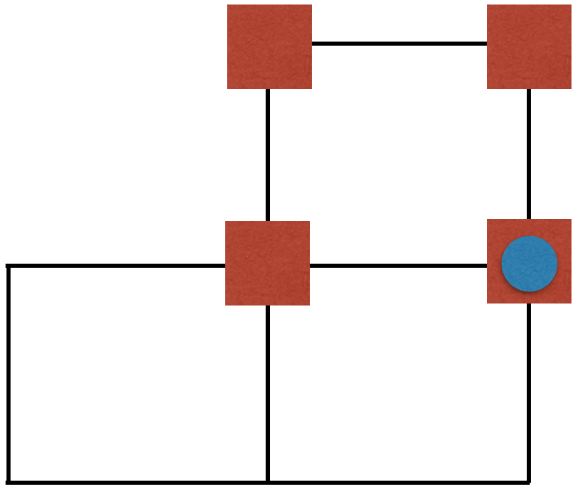


Time = 4

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

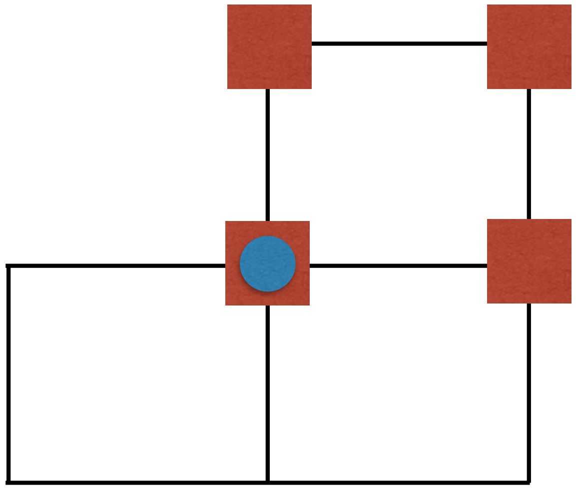


Time = 5

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

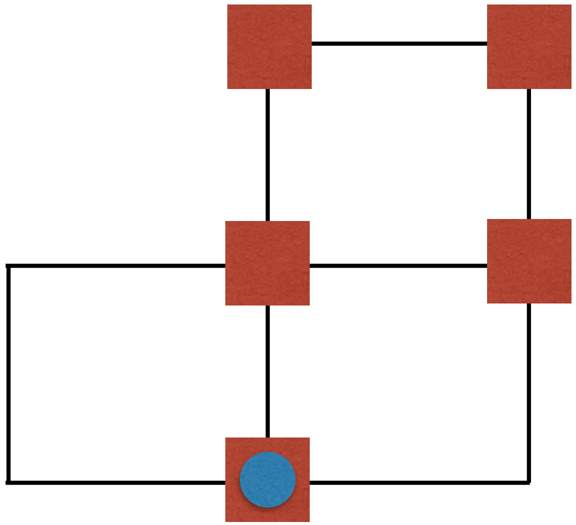


Time = 6

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

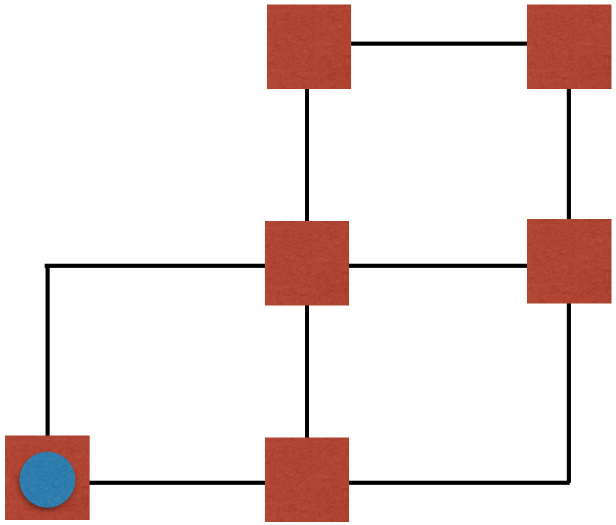


Time = 7

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

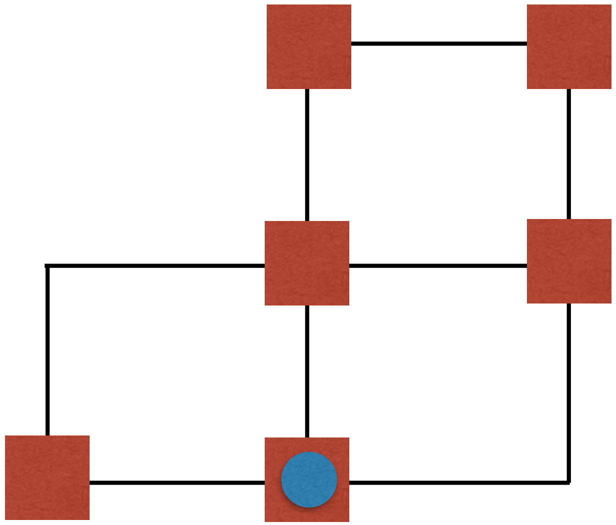


Time = 8

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

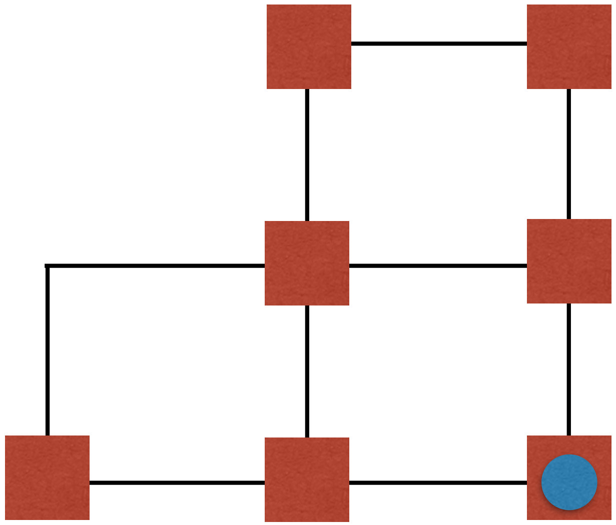


Time = 9

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

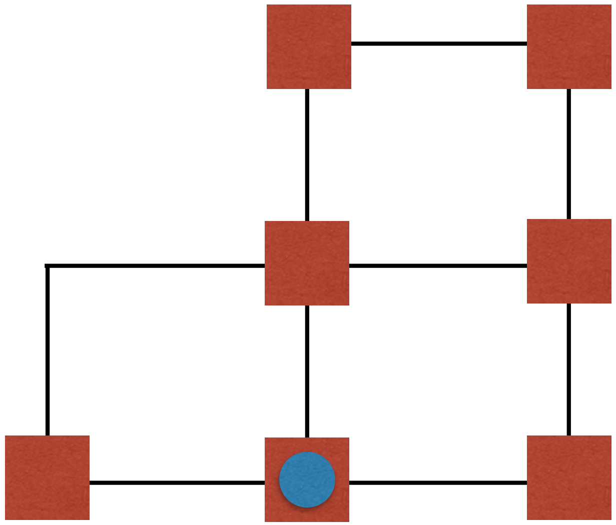


Time = 10

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

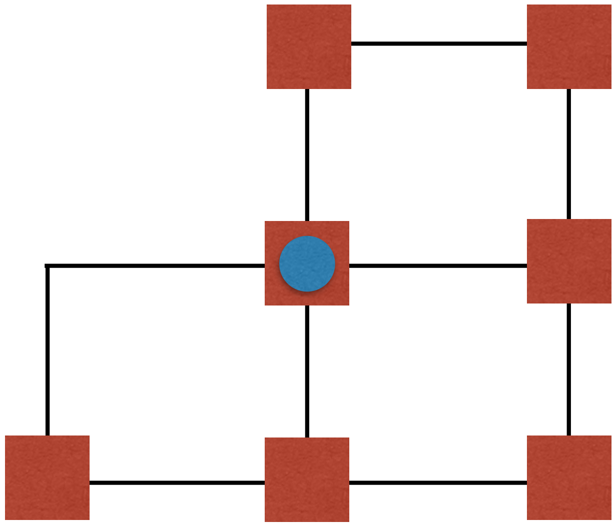


Time = 11

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

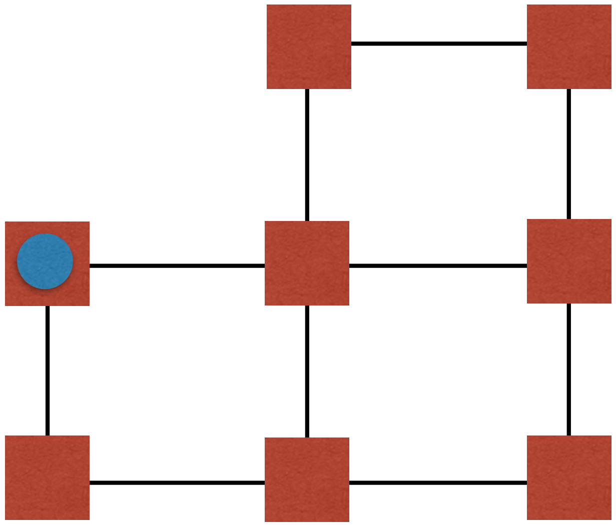


Time = 12

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.

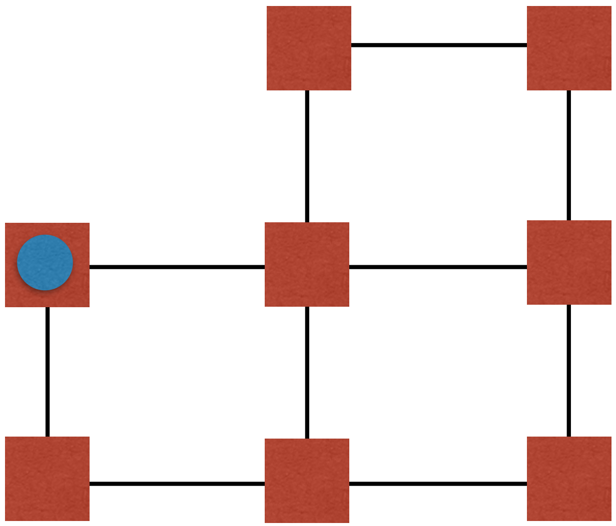


Time = 13

What is cover time ?

(for a random walk)

It is the minimum time needed by a random walk to visit all sites at least once of a confined domain.



Time = 13

Cover time = 13, for this realisation of the walk
— is a random variable

Motivation

Stochastic search processes are ubiquitous in nature:

- animals foraging for food
- proteins searching for specific DNA sequences to bind
- sperm cells searching for an oocyte to fertilize

Several of these stochastic search processes often are modeled by a single searcher performing a simple random walk.

In many situations, the search takes place in a confined domain and the targets typically are scattered over the entire domain.

Finding all these targets therefore requires an exhaustive exploration of this confined domain.

Cover time is an important observable that characterizes the efficiency of the search process.

Application in computer science

generating random spanning trees (with uniform measure) on an arbitrary connected and undirected graph G

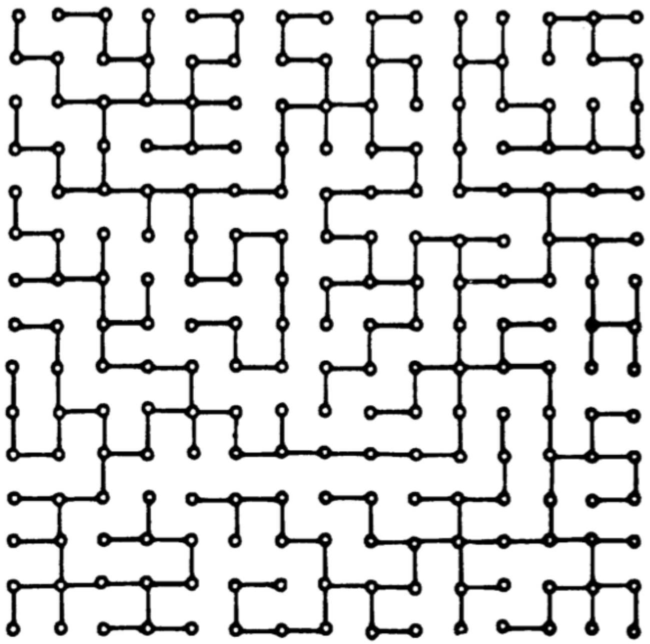


FIG. 1. A spanning tree on a 15×15 square lattice.

Manna, Dhar & Majumdar (1992)

Aldous-Broder algorithm:

- A random walk continues until all the sites are visited at least once.
- Tree is constructed by keeping the last exit bonds from each site.

Previous studies

Mostly focused on calculating the mean cover time $\langle t_c \rangle$ on regular lattices, graphs, and networks. For example,

On a complete graph with L vertices: $\langle t_c \rangle \sim L \ln L$

Aldous ('83); Broder & Karlin ('89); ….

This also applies to "rapidly mixing" walks. $t_{\mathrm{eq}} < O(L)$

e.g., regular d-dimensional lattice with $d \geq 3$.

In 2D: $\langle t_c \rangle \sim L \ln^2 L$

Brummelhuis & Hilhorst ('91); Dembo, Peres, Rosen & Zeitouni ('04); Ding ('12)

In 1D: $\langle t_c \rangle \sim L^2$

Yokoi, Hernández-Machado & Ramírez-Piscina ('90)

For distributions

It is a notoriously hard task to obtain analytical results.

Exact formal expression for cumulative distribution, on arbitrary finite graph: [Zlatanov & Kocarev, ('09)]

$$F_{\text{cover}}(t) = \sum_{i=1, i\neq z}^{n} F_{x_i}(t) - \sum_{i=1, i\neq z}^{n} \sum_{j=i+1, j\neq z}^{n} F_{x_i, x_j}(t) + \ldots + (-1)^{n-1} F_{x_1, x_2, \ldots, x_n}(t),$$

where z is the starting node of the walk.

the probability that starting from z, the RW reach for the first time one of the $x_1, x_2, \ldots, x_n$ nodes in $\leq t$ steps.

From this, it remains however extremely difficult to extract explicit results for large systems.

Rigorous result for d-dimensional regular lattice with periodic boundary conditions

Belius ('13)

For $d \geq 3:$ $z = \left[\frac{t_c - \langle t_c \rangle}{L}\right] \xrightarrow{\text{law}}$ Gumbel as $L \to \infty$

(# lattice sites)

The cumulative distribution: $F(z) = e^{-e^{-z}}$

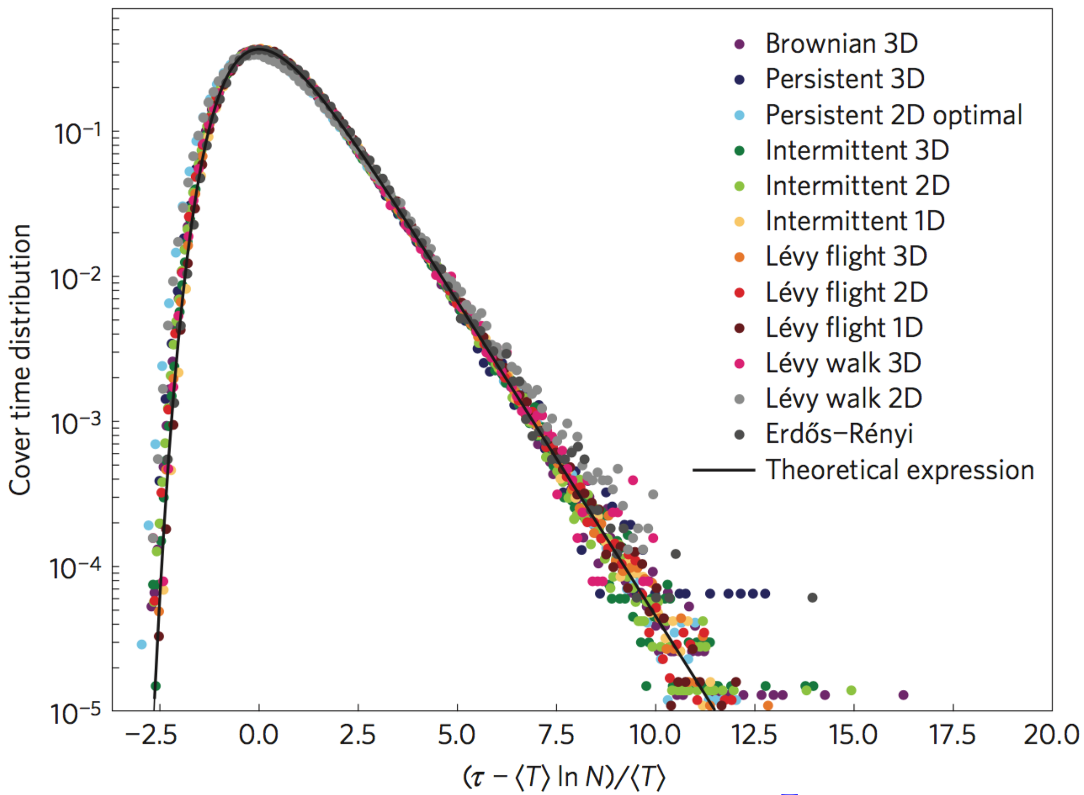
The same conclusion holds for the RW on the fully connected graph.

Turban ('15)

Transient (non-compact) Walks

The Gumbel distribution is universal for non-compact walks.

Chupeau, Bénichou & Voituriez, *Nature Physics* ('15).



Relies on the assumption that the time needed to visit a new site, from the set of unvisited sites so far, is independent of the history,

and is a Poisson process:

$$t_c = \sum_{k=1}^{L-1} t_k$$

with $p_{k,L}(t) \sim \frac{k}{\langle T\rangle} e^{-kt/\langle T\rangle}$

Exceptions

Random walks in one and two dimensions — recurrent

compact (d=1) or marginally compact (d=2)

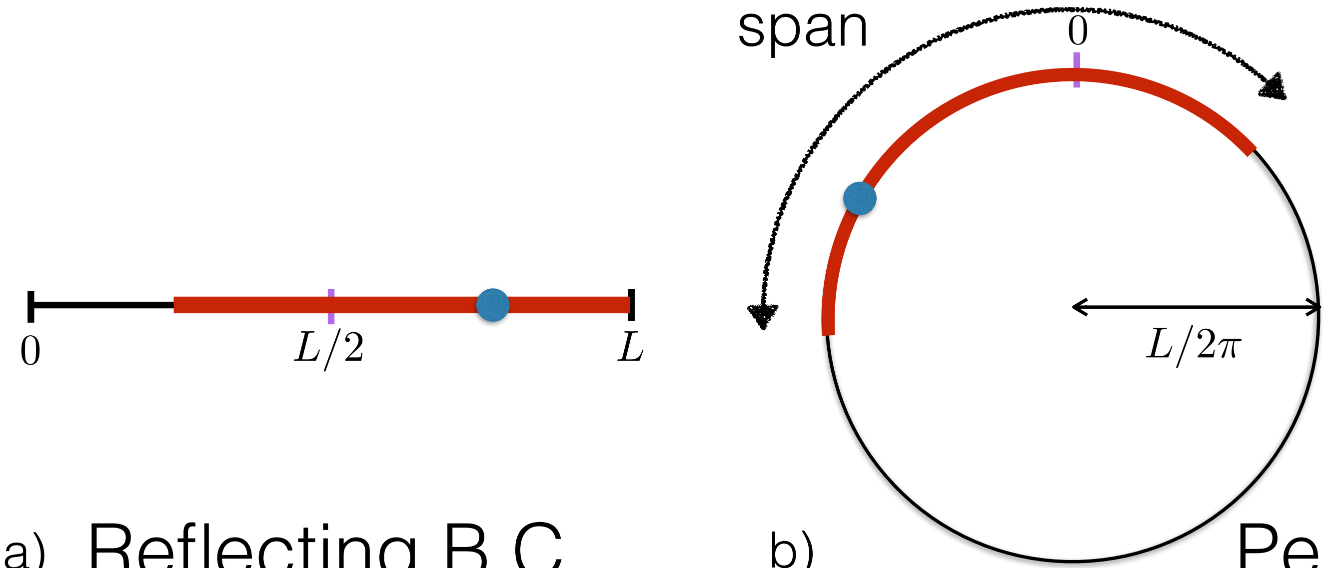
Q: Is the scaled distribution of cover time still given by a Gumbel law or is it something completely different?

In particular, for d=1.

This question is relevant for any process modeled by a one dimensional RW in a finite domain, for instance, for proteins searching for a binding site on a DNA strand.

Q: What is the role of the boundary conditions on the confined domain?

Single RW in $1d$



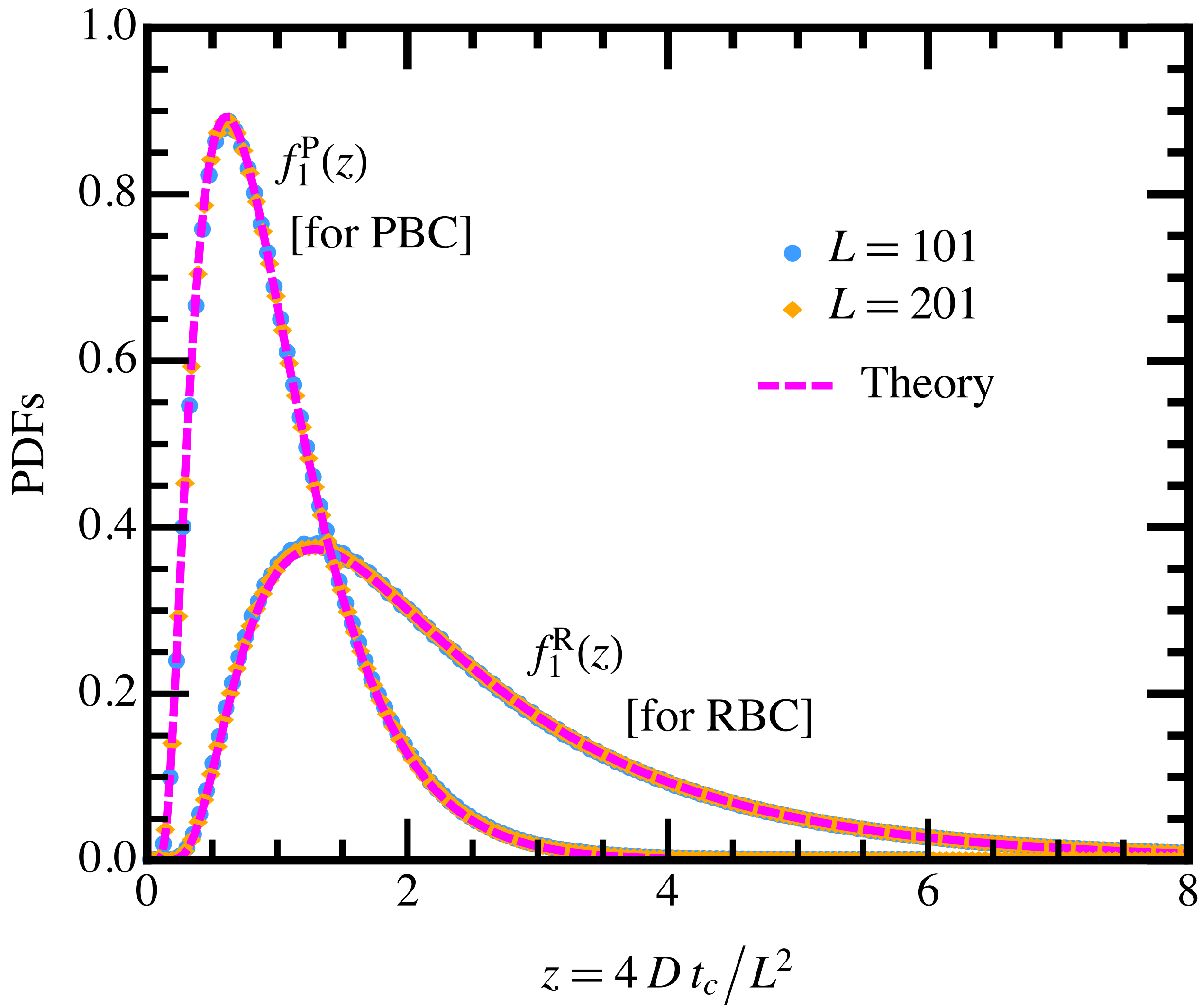
a) Reflecting B.C.

b) Periodic B.C.

$$\text{Prob.}[t_c = t|L] = \frac{4D}{L^2}\, f_1^{\mathrm{R|P}}\left(\frac{4D\,t}{L^2}\right)$$

R := RBC (a)
P := PBC (b)

Plots of the distributions of the scaled cover time



Exact expression (RBC)

$$f_1^{\mathrm{R}}(z) = \frac{1}{\sqrt{\pi} z^{3/2}} \sum_{n=0}^{\infty} \left[\sin\frac{n\pi}{2} + \cos\frac{n\pi}{2} - (-1)^n \right] (2n+1)\, e^{-(2n+1)^2/(4z)}$$

$$= \frac{\pi}{2} \sum_{n=0}^{\infty} (-1)^n (2n+1) \left[\cos[(2n+1)\pi/4]\, e^{-(2n+1)^2 \pi^2 z/16} - 2e^{-(2n+1)^2 \pi^2 z/4} \right],$$

Asymptotic behaviours

$$f_1^{\mathrm{R}}(z) \sim \begin{cases} (6/\sqrt{\pi}) z^{-3/2} e^{-9/(4z)} & \text{as } z \to 0, \\ \pi/(2\sqrt{2})\, e^{-\pi^2 z/16} & \text{as } z \to \infty. \end{cases}$$

Exact expression (PBC)

$$f_1^{\mathrm{P}}(z) = \frac{4}{\sqrt{\pi} z^{3/2}} \sum_{n=1}^{\infty} (-1)^{n+1} n^2 e^{-n^2/z}$$

$$= \sum_{n=0}^{\infty} \left[(2n+1)^2 \pi^2 z - 2 \right] e^{-\pi^2 (2n+1)^2 z/4}$$

Asymptotic behaviours

$$f_1^{\mathrm{P}}(z) \sim \begin{cases} \left(4/\sqrt{\pi}\right) z^{-3/2} e^{-1/z} & \text{as } z \to 0, \\ \pi^2 z\, e^{-\pi^2 z/4} & \text{as } z \to \infty. \end{cases}$$

For N independent RWs

(all starting at the same site)

The cover time is the minimum time needed for all sites to be visited at least once by at least one of the walkers.

This problem naturally arises in various search problems, where there is a team of N independent searchers, as opposed to a single searcher.

For N=2, with the PBC: $\langle t_c \rangle \simeq \frac{1}{4} L^2$ $\left[\simeq \frac{1}{2} L^2 \text{ for } N = 1 \right]$

Hemmer and S. Hemmer ('98)

Q: How does it decrease for large N? (also distributions)

The mean

$$\frac{4D\,\langle t_c \rangle}{L^2} \approx \frac{1}{4\ln N}$$ for both RBC & PBC

The fluctuations around the mean

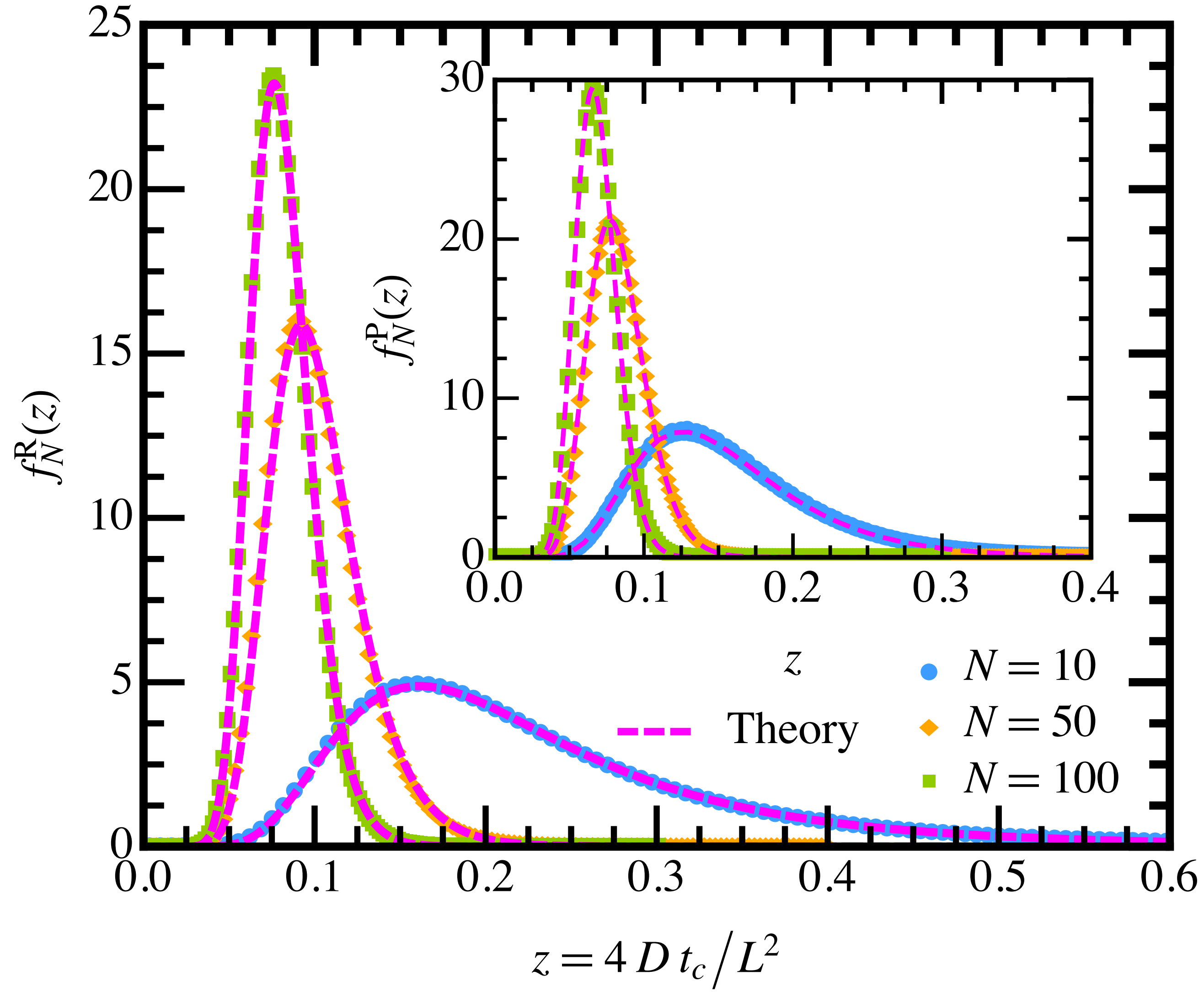
$$\frac{4D\,t_c}{L^2} \approx \frac{1}{4\ln N} - \frac{1}{4(\ln N)^2}\,\chi_{\mathrm{R,P}}$$

are sensitive to the boundary conditions.

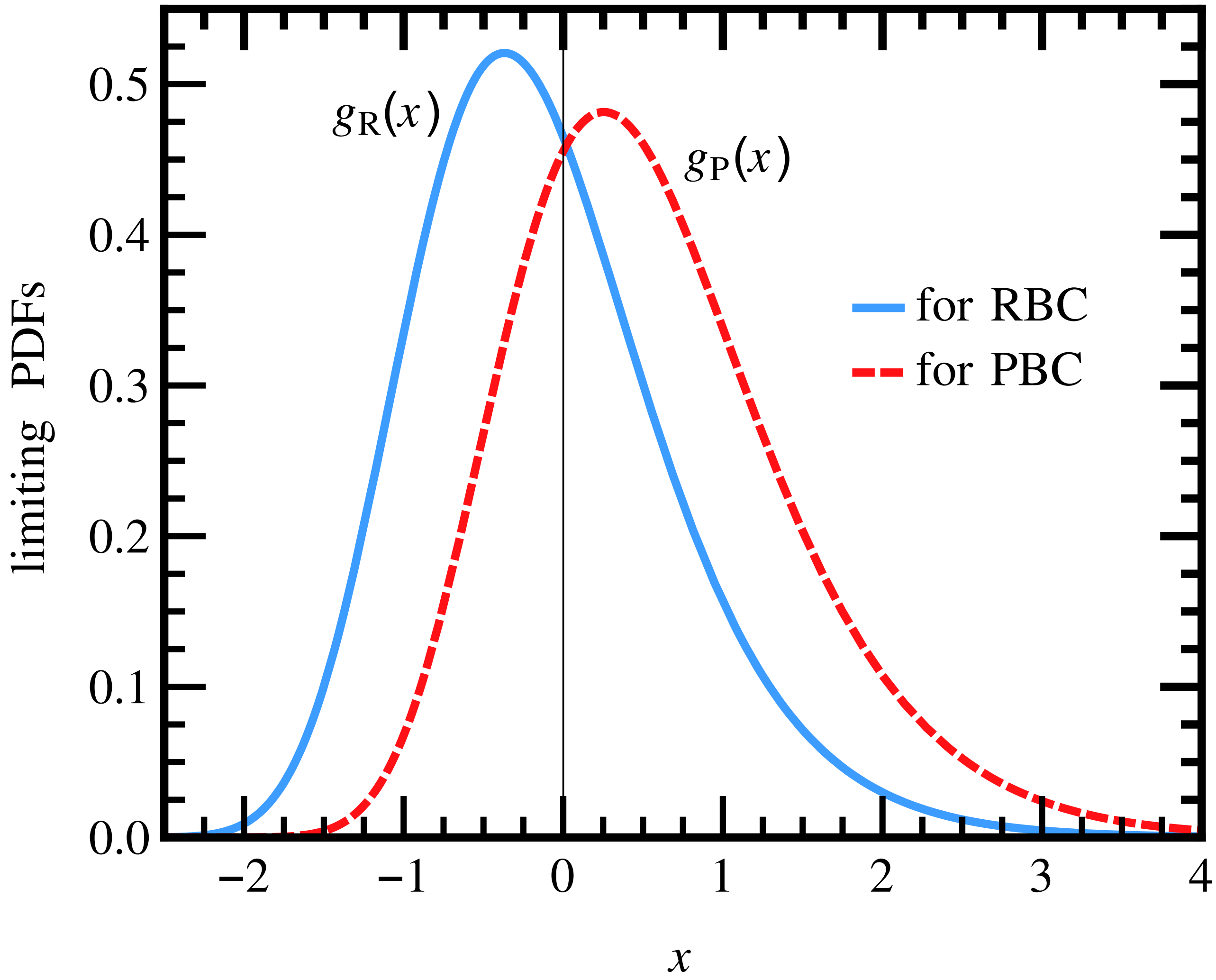
RBC: $\mathrm{Prob.}[\chi_{\mathrm{R}} = x] = g_{\mathrm{R}}(x) = 2\,e^{-x-e^{-x}}\left(1 - e^{-e^{-x}}\right)$

PBC: $\mathrm{Prob.}[\chi_{\mathrm{P}} = x] = g_{\mathrm{P}}(x) = 4\,e^{-2x}\,K_0(2\,e^{-x})$

Plots of the distributions of the scaled cover time



Plot of limiting PDFs



Summary

- We have obtained the full PDF of the cover time for N independent Brownian motions in one dimension, both for RBC and for PBC.

  Previously, only the mean was known in 1D for $N=1$ and $N=2$.

- This results provide an instance of exact cover time distributions for recurrent random walks, demonstrating clearly that this is different from a Gumbel law found recently for transient (i.e., nonrecurrent) walks.

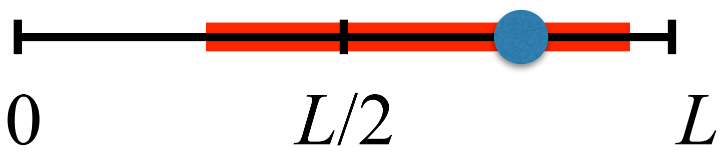
- In addition, we have shown that in the limit of large N, the cover time approaches its average value $\approx L^2/(16D \ln N)$ with fluctuations decaying as $1/(\ln N)^2$.

  The centered and scaled distributions converge to two distinct and nontrivial N independent scaling functions for RBC and PBC.

- Open problem: determining the full PDF in $d = 2$ for one or multiple walkers (another instance of recurrent RW) — mean has been studied.

Methods

Single walker (reflecting case)



The cover time = The first time when the walker has hit both boundaries at $x = 0$ and $x = L$.

Prob.[$t_c > t | x_0$] = Prob.[at least one of the boundaries has not been hit up to time t]

= Prob.[0 is unhit up to time t] → $S_{AR}(x_0, t)$

\+ Prob.[L is unhit up to time t] → $S_{RA}(x_0, t)$

– Prob.[both are unhit up to time t]. → $S_{AA}(x_0, t)$

$$\frac{\partial S(x_0, t)}{\partial t} = D \frac{\partial^2 S(x_0, t)}{\partial x_0^2}$$

← satisfy

N independent walkers (reflecting case)

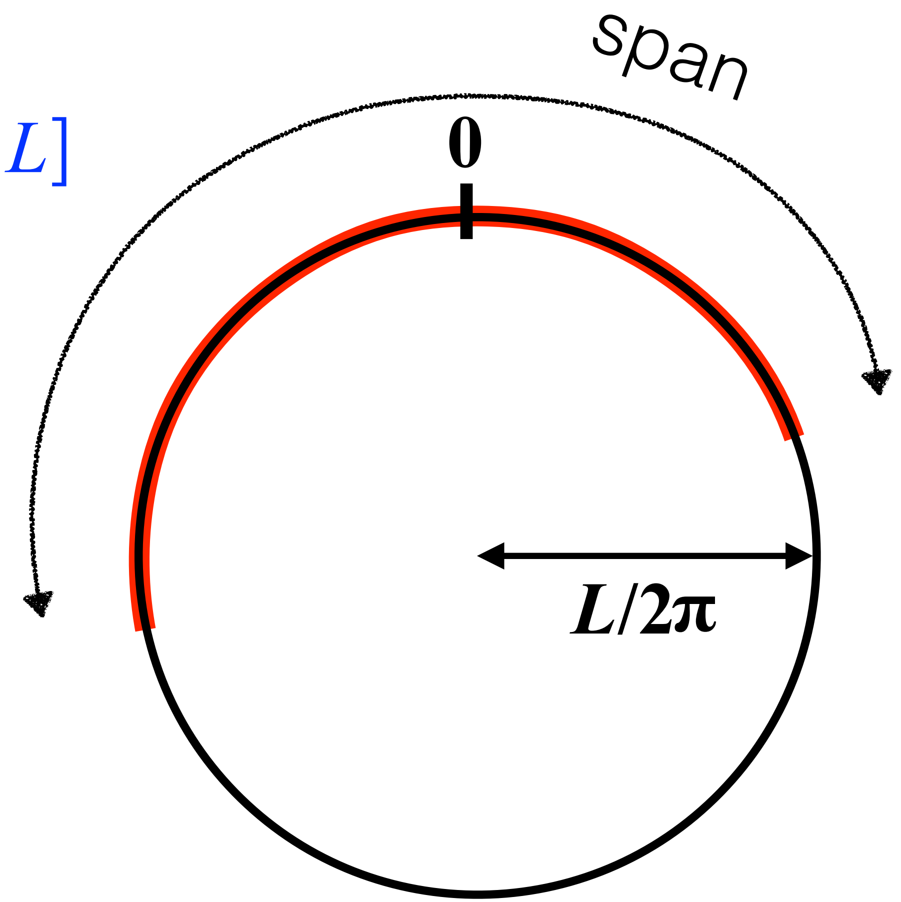
$$\text{Prob.}[t_c > t | x_0, N] = [S_{\text{AR}}(x_0, t)]^N + [S_{\text{RA}}(x_0, t)]^N - [S_{\text{AA}}(x_0, t)]^N$$

Periodic case

$\text{Prob.}[t_c > t|L] = \text{Prob.}[\text{span}(t) < L]$

the length of the
covered region
by the walker up to time t.

(on the infinite line)



Since the ring has not been traversed fully at time t,
the walker does not realize that it is on a ring.

PDF(span,t) is known.
(single walker: textbook problem)

For N walkers:
Kundu, Majumdar & Schehr, ('13)