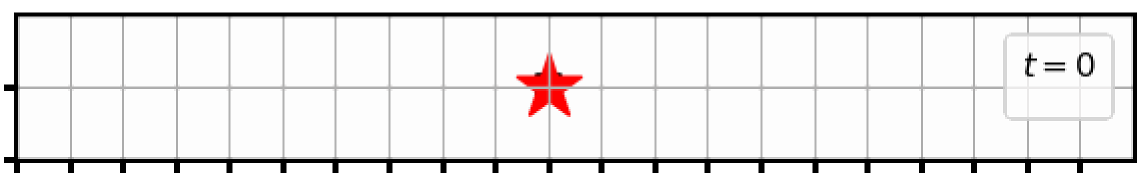
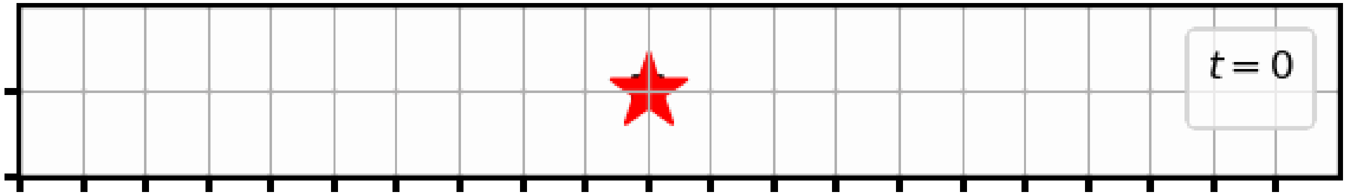
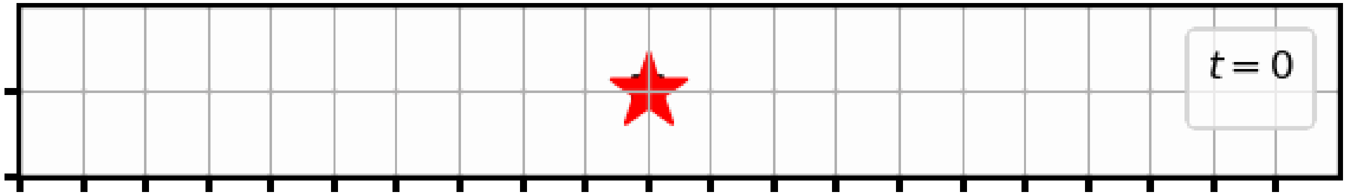
Visitation statistics of 1d random walks

Léo Régnier
(with M. Dolgushev, S. Redner and O. Bénichou)



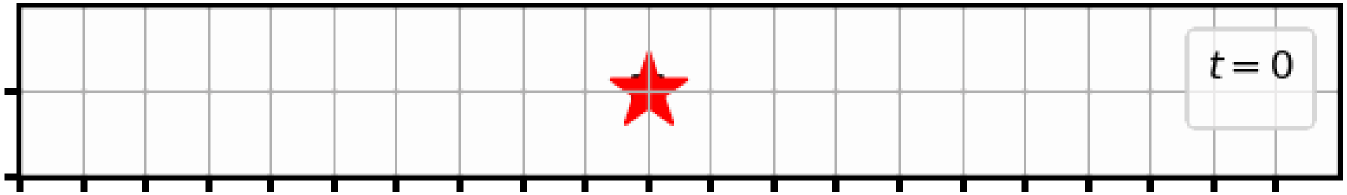


Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.



Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.

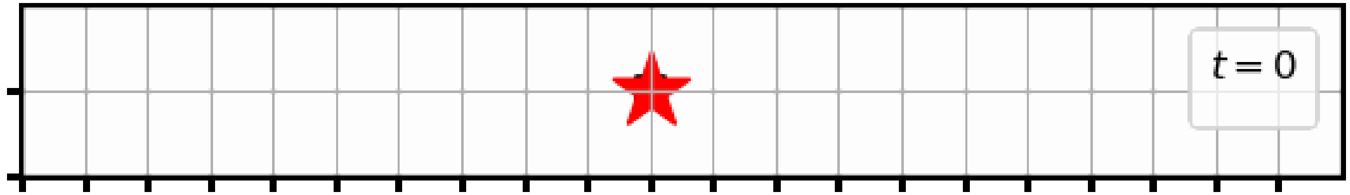
$$X_t = \ell_1 + \ldots + \ell_t \qquad \ell_i = \pm 1$$



Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.

$$X_t = \ell_1 + \ldots + \ell_t \qquad \ell_i = \pm 1$$

The jumps distribution is...



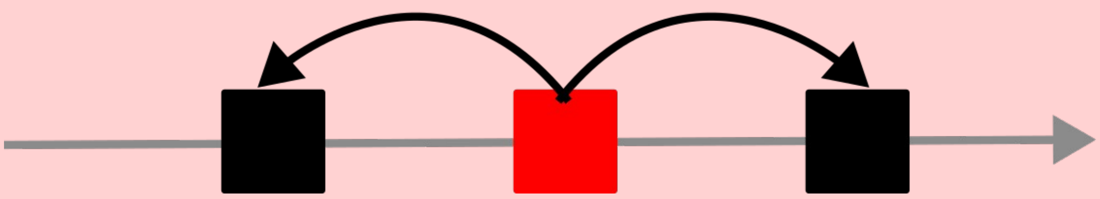
Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.

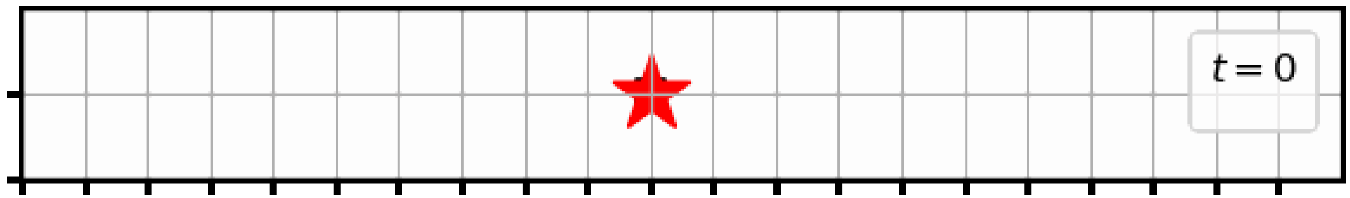
$$X_t = \ell_1 + \ldots + \ell_t \qquad \ell_i = \pm 1$$

The jumps distribution is...

- **Symmetric,**

$p(\ell_i = 1) = 1/2$



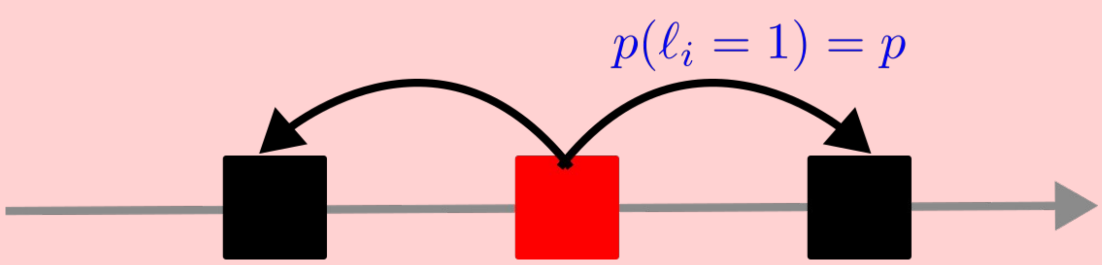


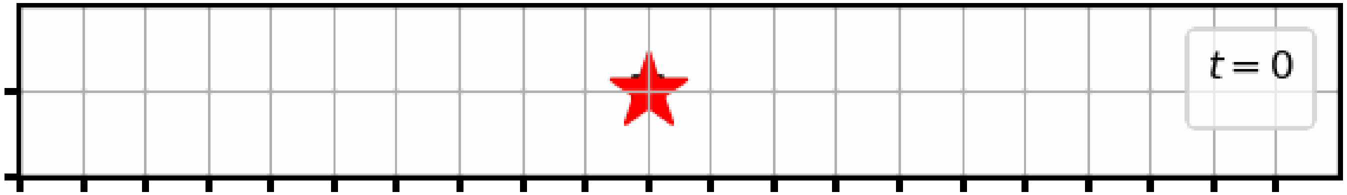
Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.

$$X_t = \ell_1 + \ldots + \ell_t \qquad \ell_i = \pm 1$$

The jumps distribution is...

- Symmetric,
- **Asymmetric,**



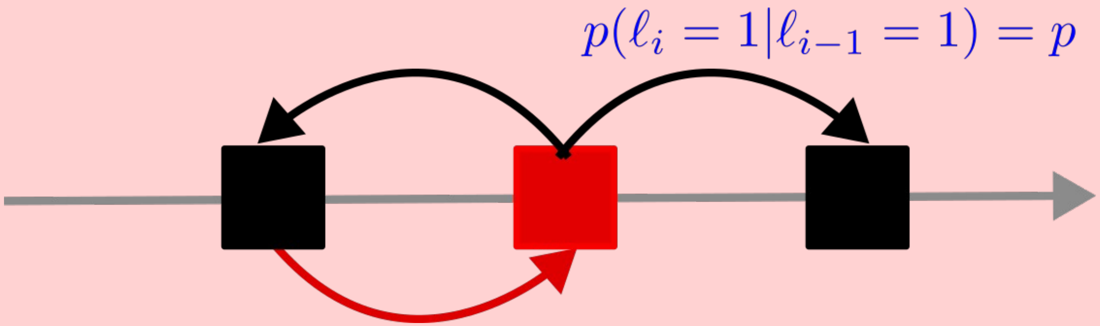


Model: 1d lattice, a random walker does (random) jumps on adjacent sites every time unit.

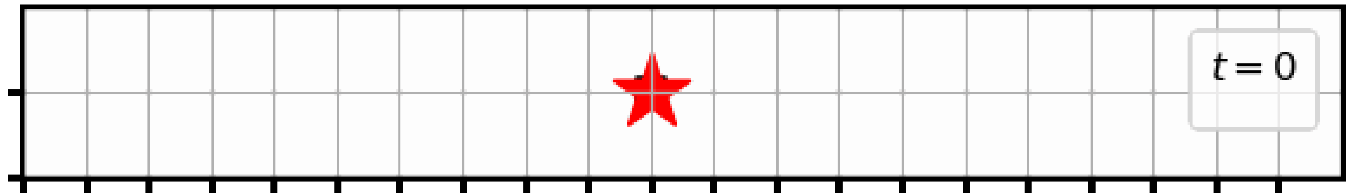
$$X_t = \ell_1 + \ldots + \ell_t \qquad \ell_i = \pm 1$$

The jumps distribution is...

- Symmetric,
- Asymmetric,
- **Correlated,**

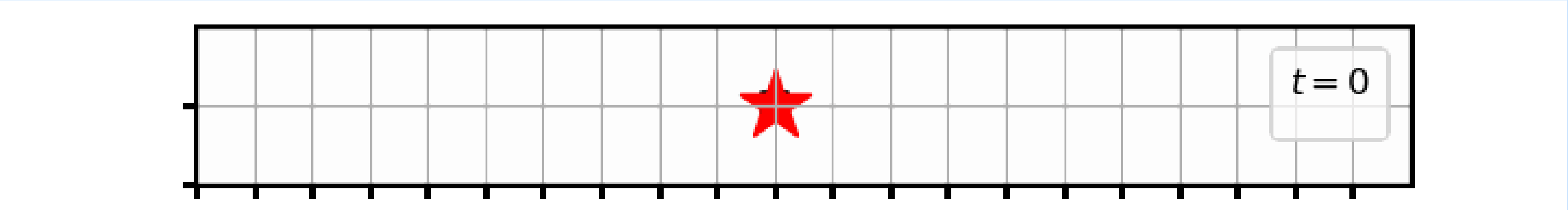


The random walk explores the 1d line up to a time t.



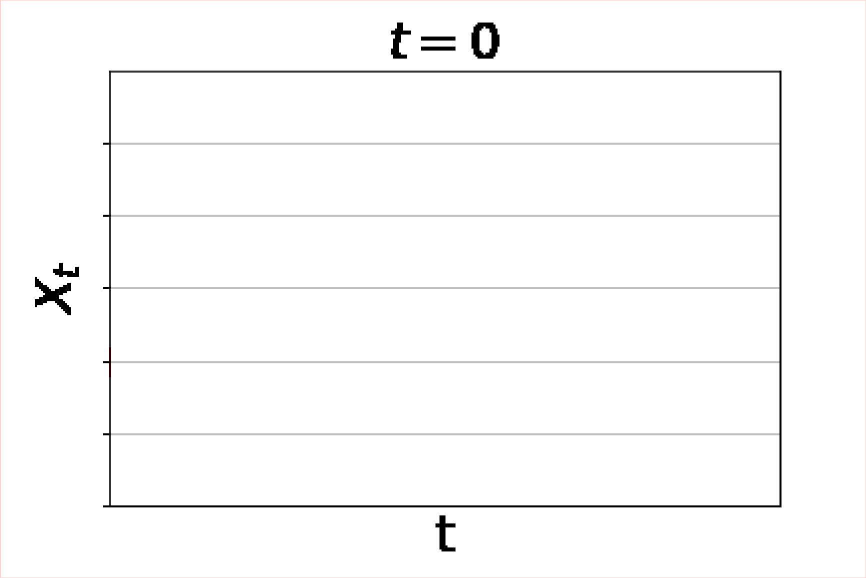
How do we characterize the territory explored by the random walk?

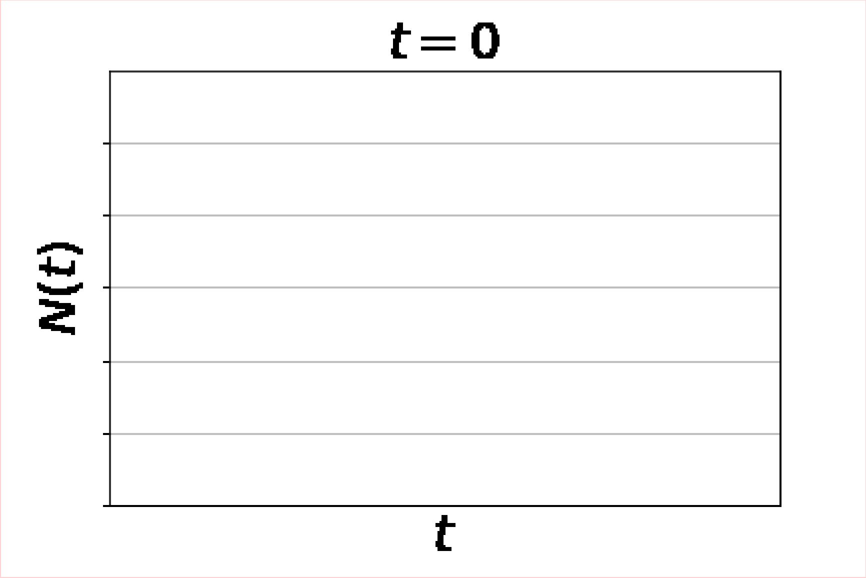
The random walk explores the 1d line up to a time t.



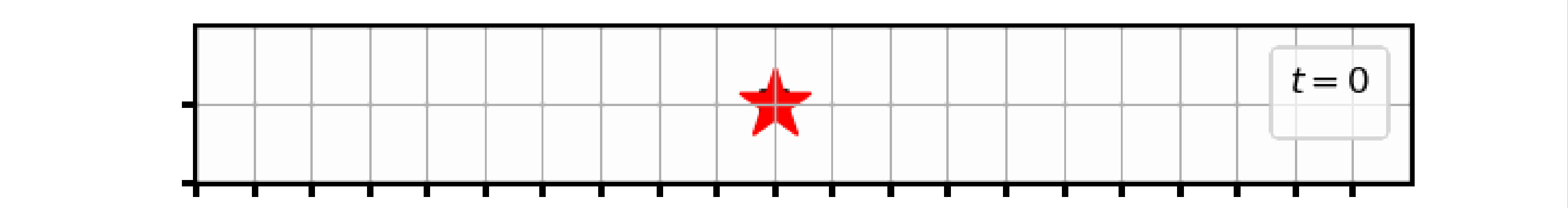
How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$





The random walk explores the 1d line up to a time t.

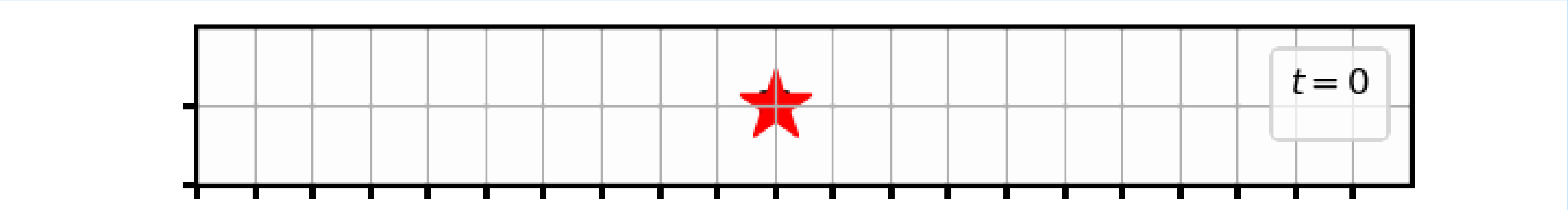


How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (Vineyard 1963, Annesi 2019, Wiese 2020):

The random walk explores the 1d line up to a time t.



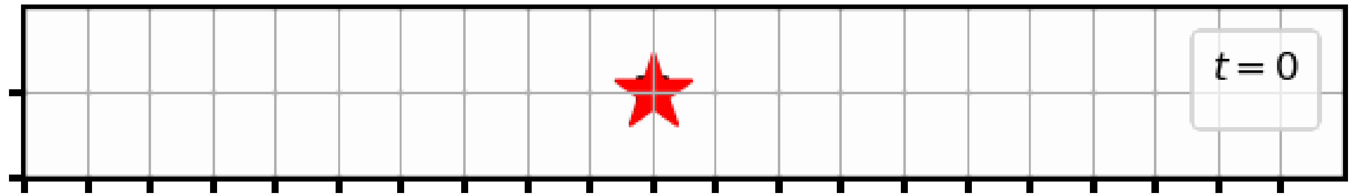
How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (symmetric case):

- The average $\langle N(t) \rangle \sim \sqrt{\frac{8t}{\pi}}$

The random walk explores the 1d line up to a time t.



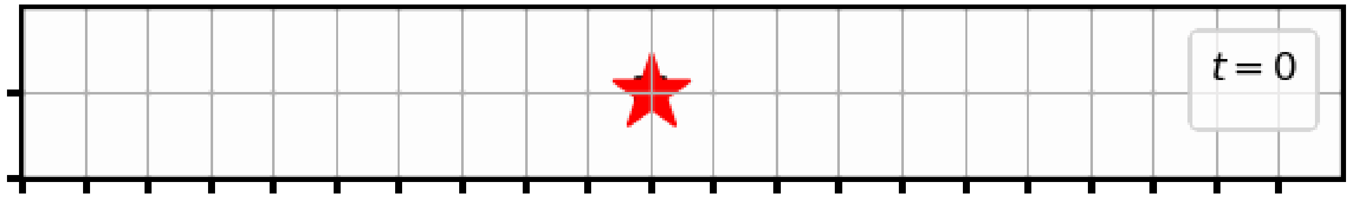
How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (symmetric case):

- The average $\langle N(t) \rangle \sim \sqrt{\frac{8t}{\pi}}$
- The variance $\langle N(t)^2 \rangle = 4\ln(2)\, t$

The random walk explores the 1d line up to a time t.



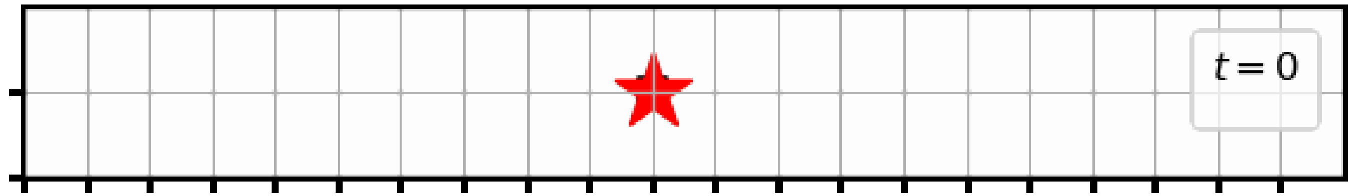
How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (symmetric case):

- The average $\langle N(t)\rangle \sim \sqrt{\frac{8t}{\pi}}$
- The variance $\left\langle N(t)^2\right\rangle = 4\ln(2)\,t$
- The distribution $\mathbb{P}(N(t) \geq n)$

The random walk explores the 1d line up to a time t.



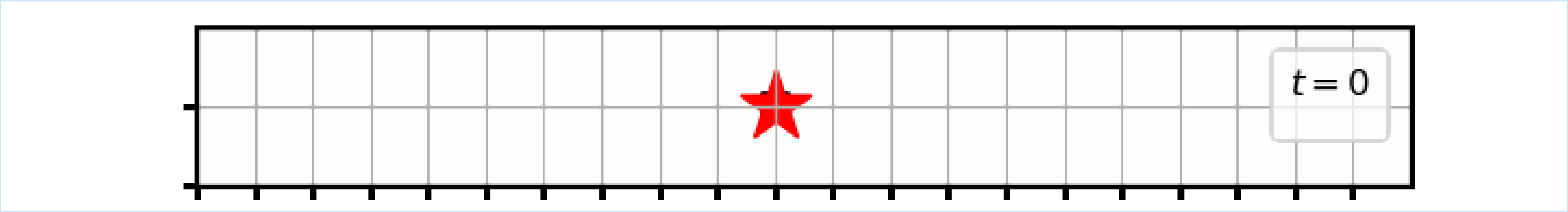
How do we characterize the territory explored by the random walk?

By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (symmetric case):

- The average $\langle N(t)\rangle \sim \sqrt{\frac{8t}{\pi}}$
- The variance $\langle N(t)^2\rangle = 4\ln(2)\,t$
- The distribution $\mathbb{P}(N(t) \geq n)$
- The covariance $\mathrm{Cov}(N(t_1), N(t_2))$

The random walk explores the 1d line up to a time t.



How do we characterize the territory explored by the random walk?

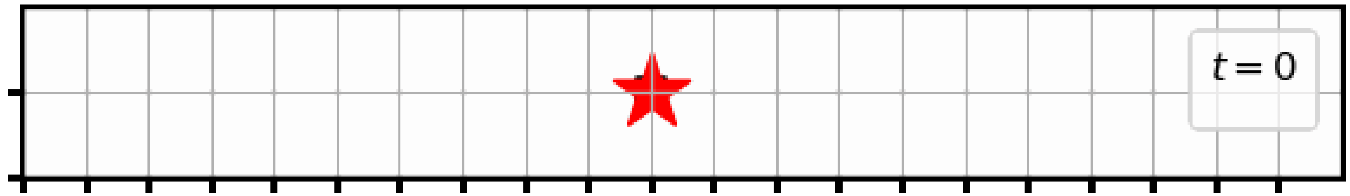
By the number of distinct sites visited, $N(t)$

Several known results can be found in the litterature (symmetric case):

- The average $\langle N(t) \rangle \sim \sqrt{\frac{8t}{\pi}}$
- The variance $\langle N(t)^2 \rangle = 4 \ln(2)\, t$
- The distribution $\mathbb{P}(N(t) \geq n)$
- The covariance $\mathrm{Cov}(N(t_1), N(t_2))$

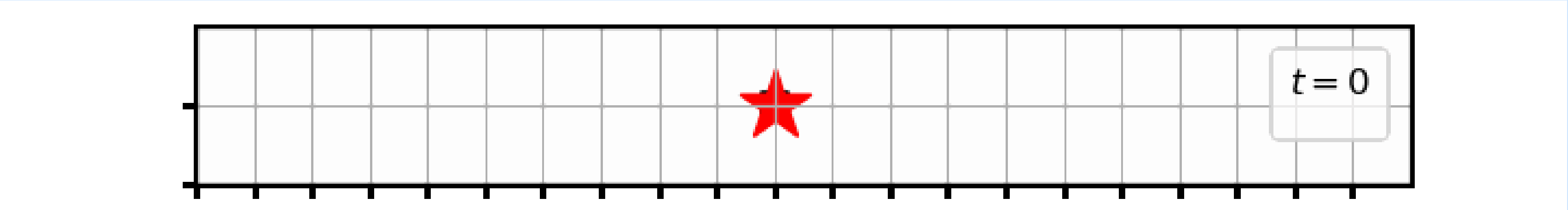
Partial characterization of the exploration process!

The random walk explores the 1d line up to a time t.



How do we fully characterize the territory explored by the random walk?

The random walk explores the 1d line up to a time t.

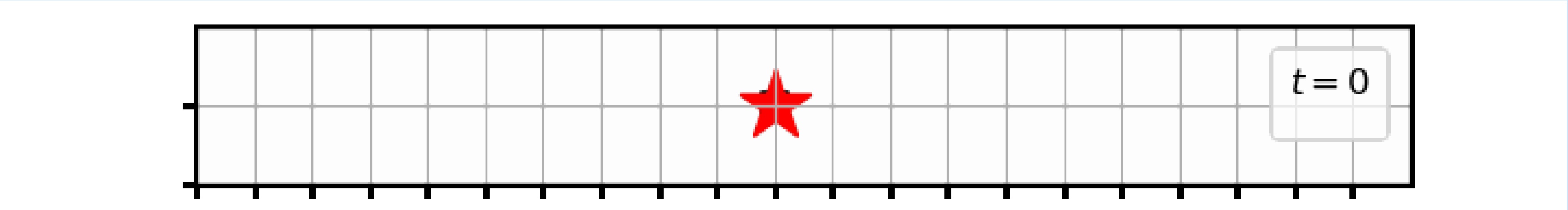


How do we fully characterize the territory explored by the random walk?

By its muliple time points distribution...

$$\mathbb{P}(N(t_1) = n_1, \ldots, N(t_k) = n_k)$$

The random walk explores the 1d line up to a time t.



How do we fully characterize the territory explored by the random walk?

By its muliple time points distribution...

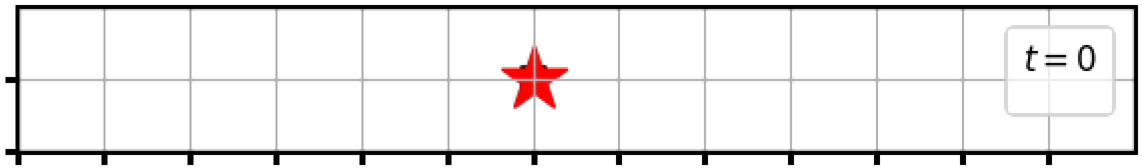
$$\mathbb{P}(N(t_1) = n_1, \ldots, N(t_k) = n_k)$$

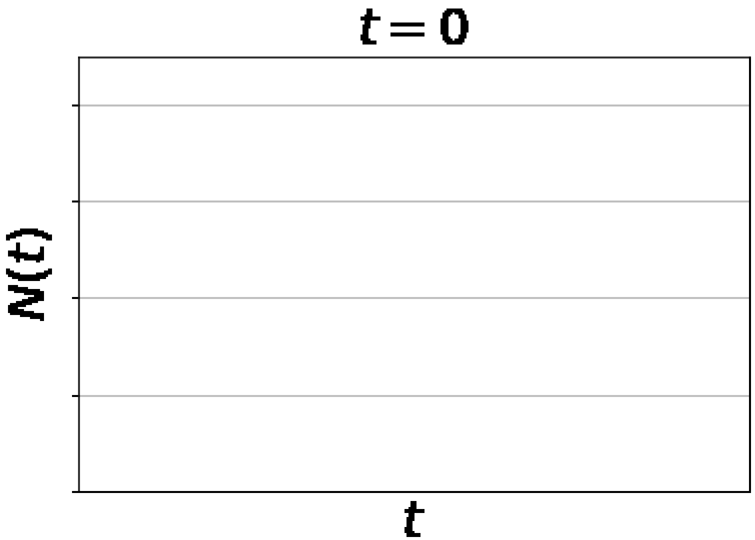
... but how to obtain it?

We intoduce new tools!

We intoduce new tools!

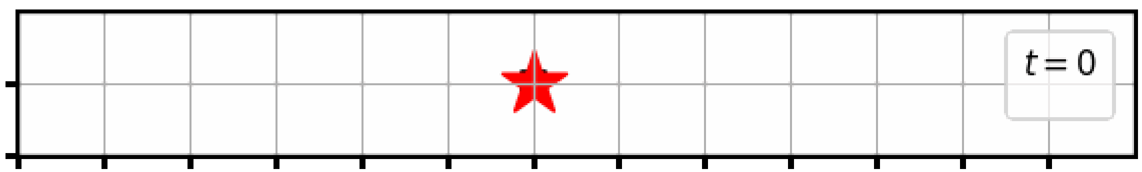
- τ_n : time elapsed between visit of site n and n+1

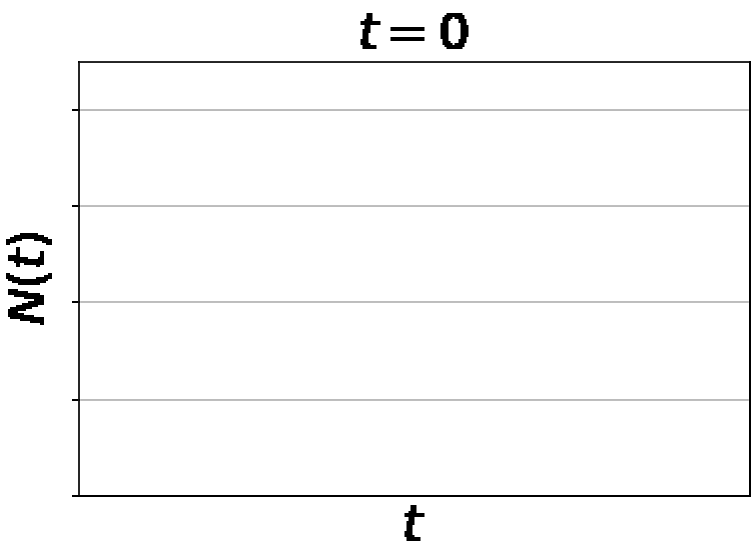




We intoduce new tools!

- τ_n : time elapsed between visit of site n and n+1

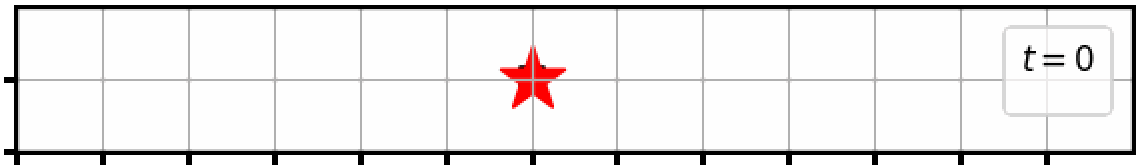


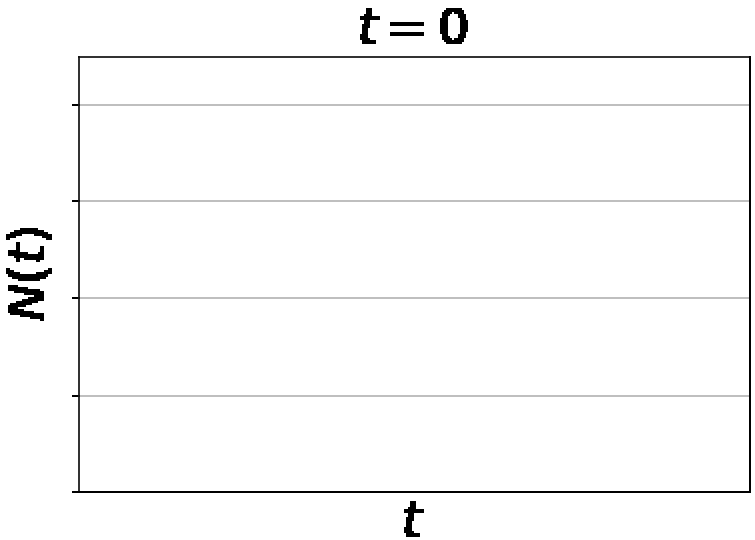


But why?

We intoduce new tools!

- τ_n : time elapsed between visit of site n and n+1



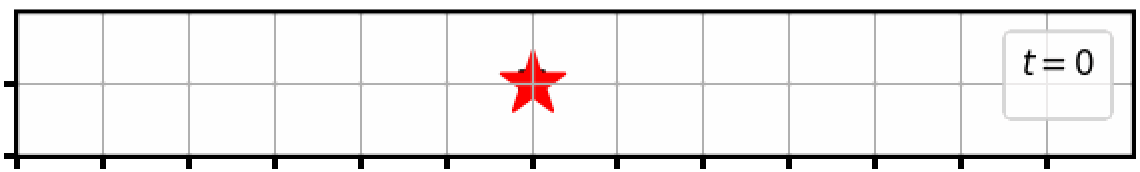


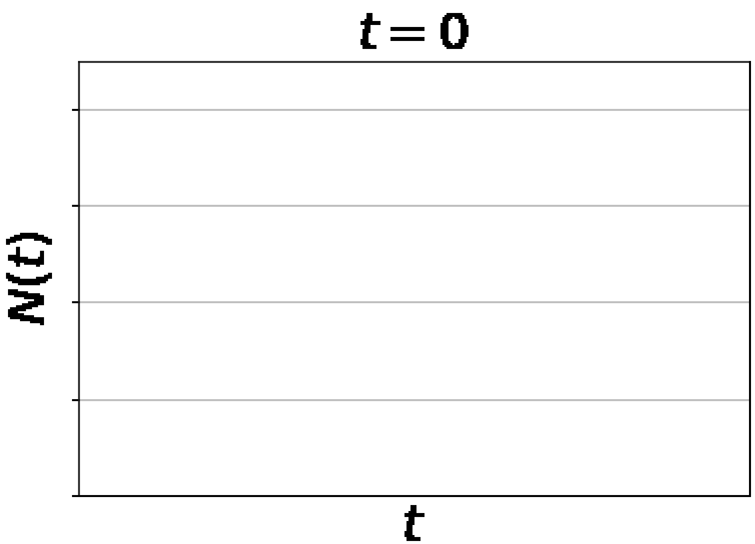
But why?

Because they verify the following key properties,

Having visited at least n sites at t is the same as having visited n sites before t, We intodice new tools!

- τ_n: time elapsed between visit of site n and n+1





But why?

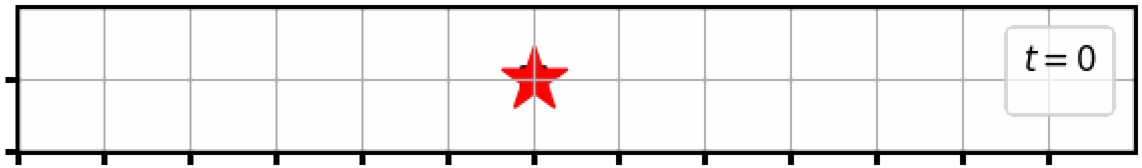
Because they verify the following key properties,

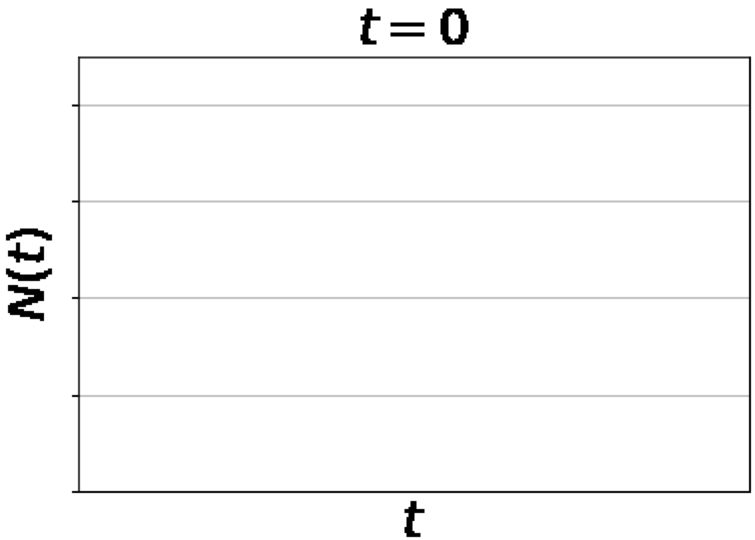
Having visited at least n sites at t is the same as having visited n sites before t,

$$\{N(t) \geq n\} = \{\tau_0 + \ldots + \ldots \tau_{n-1} \leq t\}$$

We intoduce new tools!

- τ_n : time elapsed between visit of site n and n+1





But why?

Because they verify the following key properties,

Having visited at least n sites at t is the same as having visited n sites before t,

$$\{N(t) \geq n\} = \{\tau_0 + \ldots + \ldots \tau_{n-1} \leq t\}$$

They are independent (up to a conditionning on the positions for the biased walks)

In the following, we detail the general method to obtain the multiple time distribution,

$$\mathbb{P}(N(t_1) = n_1, \ldots, N(t_k) = n_k)$$

In the following, we detail the general method to obtain the multiple time distribution,

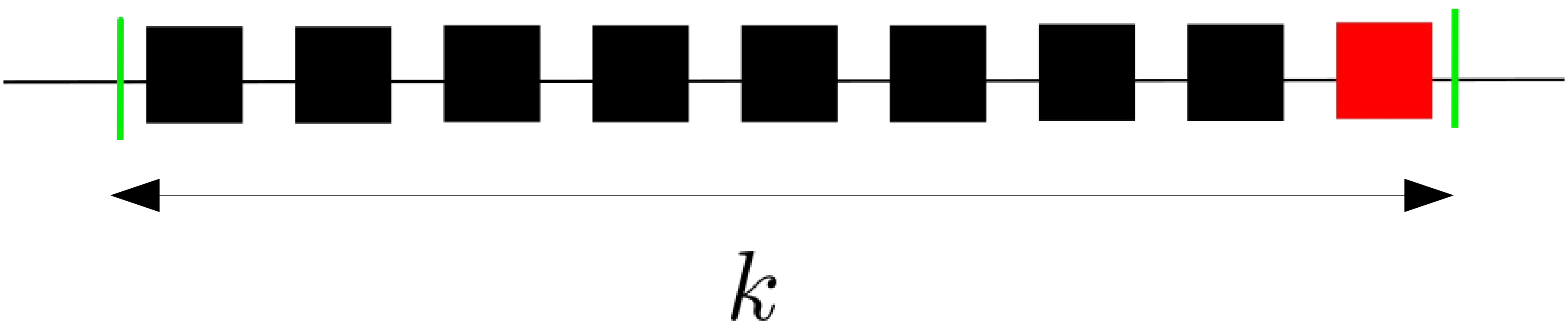
$$\mathbb{P}(N(t_1) = n_1, \ldots, N(t_k) = n_k)$$

We start by giving the general steps to obtain the **cumulative two times distribution**, for a symmetric walk,

$$\mathbb{P}(N(t_1) \geq n_1; N(t_2) \geq n_2)$$

Step 1: We start by solving the exit time problem from an interval:

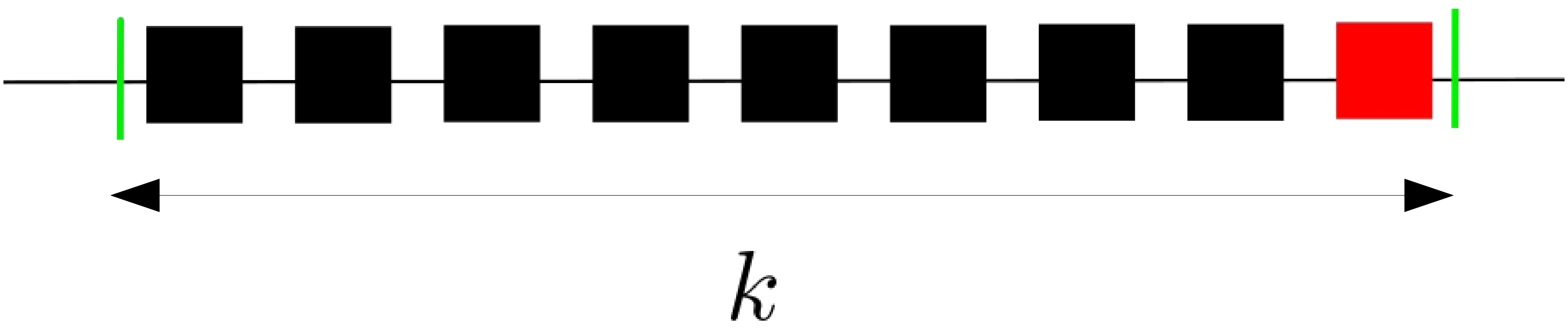
Once a new site is visited, we start at one end of the visited domain of length k.



Step 1: We start by solving the exit time problem from an interval:

Once a new site is visited, we start at one end of the visited domain of length k.
The first exit time is τ_k.

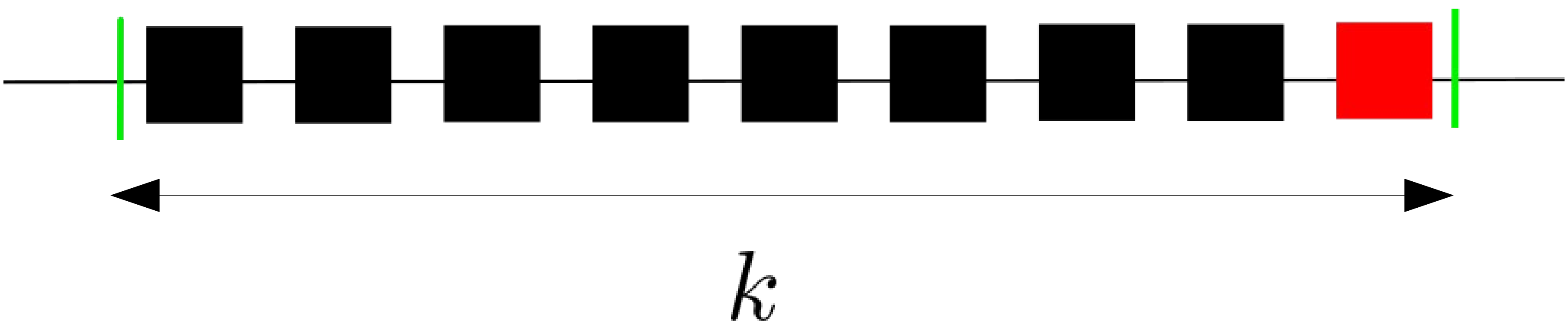
$$\boxed{\mathbb{P}(\tau_k = t) = F(t,k)}$$



Step 1: We start by solving the exit time problem from an interval:

Once a new site is visited, we start at one end of the visited domain of length k. The first exit time is τ_k.

$$\boxed{\mathbb{P}(\tau_k = t) = F(t, k)}$$

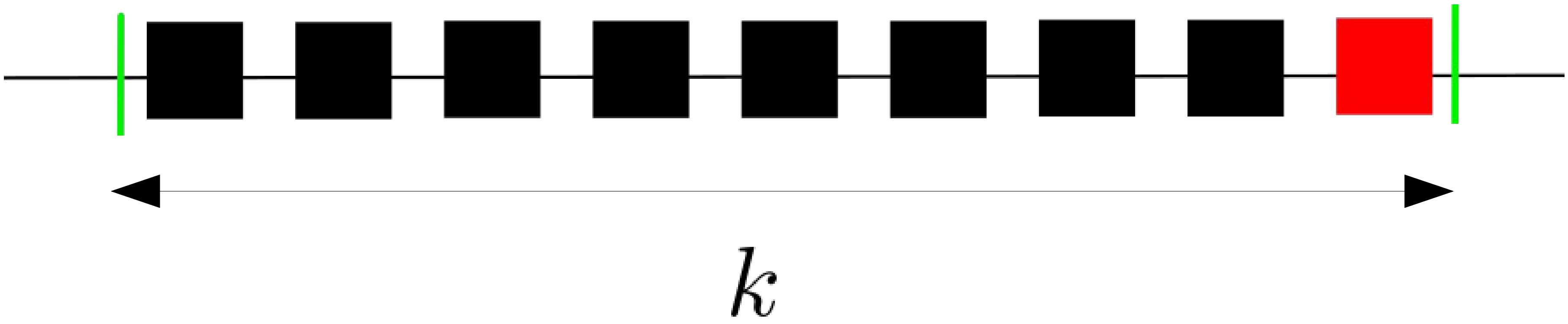


All calculations are performed in the Laplace domain (in time) $\mathcal{L}\{f(t)\} \equiv \hat{f}(s) \equiv \sum_{t=0}^{\infty} f(t)e^{-st}$

Step 1: We start by solving the exit time problem from an interval:

Once a new site is visited, we start at one end of the visited domain of length k. The first exit time is τ_k.

$$\mathbb{P}(\tau_k = t) = F(t,k)$$



All calculations are performed in the Laplace domain (in time) $\mathcal{L}\{f(t)\} \equiv \hat{f}(s) \equiv \sum_{t=0}^{\infty} f(t) e^{-st}$

EX: Symmetric random walk $\widehat{F}(s,k) = \frac{\sinh(\sqrt{2s}(k-1)) + \sinh(\sqrt{2s})}{\sinh(\sqrt{2s}k)}$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation,

$$\mathbb{P}\big(N(t_1) \geq n_1; N(t_2) \geq n_2\big) = \mathbb{P}\big(\tau_0 + ... + \tau_{n_1-1} \leqslant t_1; \tau_0 + ... + \tau_{n_2-1} \leqslant t_2\big)$$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

$$\mathbb{P}\big(N(t_1) \geq n_1; N(t_2) \geq n_2\big) = \mathbb{P}\big(\tau_0 + ... + \tau_{n_1-1} \leqslant t_1; \tau_0 + ... + \tau_{n_2-1} \leqslant t_2\big)$$

$$\begin{aligned}\mathcal{L}\{\mathbb{P}(N(t_1) \geq n_1, N(t_2) \geq n_2)\} &= \frac{1}{s_1 s_2}\widehat{F}(s_1+s_2,0)\dots\widehat{F}(s_1+s_2,n_1-1)\widehat{F}(s_2,n_1)\dots\widehat{F}(s_2,n_2-1)\\ &= \frac{1}{s_1 s_2}\exp\left(\sum_{k<n_1}\ln\widehat{F}(s_1+s_2,k) + \sum_{n_1\leq k<n_2}\ln\widehat{F}(s_2,k)\right)\end{aligned}$$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

Step 3: We take the continuum limit.

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

Step 3: We take the continuum limit.

We obtain in this limit $\widehat{F}(s,k) = 1 + g(s,k)$ $\quad h(s,n) = \exp \int_n^0 g(s,k)dk$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

Step 3: We take the continuum limit.

We obtain in this limit $\widehat{F}(s,k) = 1 + g(s,k)$ $\quad h(s,n) = \exp \int_n^0 g(s,k)dk$

$$\mathcal{L}\{\mathbb{P}(N(t_1) \geq n_1, N(t_2) \geq n_2)\} = \frac{h(s_2 n_1)}{s_1 s_2 h(s_1+s_2, n_1) h(s_2, n_2)}$$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

Step 3: We take the continuum limit.

We obtain in this limit $\widehat{F}(s,k) = 1 + g(s,k)$ $\quad h(s,n) = \exp \int_n^0 g(s,k)dk$

$$\mathcal{L}\{\mathbb{P}(N(t_1) \geq n_1, N(t_2) \geq n_2)\} = \frac{h(s_2 n_1)}{s_1 s_2 h(s_1+s_2, n_1) h(s_2, n_2)}$$

EX: Symmetric random walk

$$\widehat{F}(s,k) = 1 - \sqrt{2s}\tanh\left(\sqrt{sk^2/2}\right) \qquad h(s,n) = \cosh^2\left(\sqrt{sn^2/2}\right)$$

Step 1: We start by solving the exit time problem from an interval.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k 's.

Step 3: We take the continuum limit.

The generalization to the **k times distribution** is now straightforward, following the same three steps, valid for symmetric and correlated jumps,

$$\mathcal{L}\left\{\mathbb{P}\left(N(t_1) \geqslant n_1; \ldots; N(t_k) \geqslant n_k\right)\right\}$$

$$= \frac{1}{s_1 \ldots s_k} \frac{1}{h(s_1 + \ldots + s_k, n_1)} \frac{h(s_2 + \ldots + s_k, n_1)}{h(s_2 + \ldots + s_k, n_2)} \cdots \frac{h(s_k, n_{k-1})}{h(s_k, n_k)}$$

Does the knowledge of visited territory at early times influence the one at long times?

Does the knowledge of visited territory at early times influence the one at long times?

Indeed, it does: There are long range correlations

$$\mathbb{P}(N(t_1)=n_1, N(t_2)=n_2) - \mathbb{P}(N(t_1)=n_1)\mathbb{P}(N(t_2)=n_2) = C_{a_1,a_2}\frac{t_1^{1/2}}{t_2^{3/2}}$$

Does the knowledge of visited territory at early times influence the one at long times?

Indeed, it does: There are long range correlations

$$\mathbb{P}(N(t_1)=n_1, N(t_2)=n_2) - \mathbb{P}(N(t_1)=n_1)\mathbb{P}(N(t_2)=n_2) = C_{a_1,a_2}\frac{t_1^{1/2}}{t_2^{3/2}}$$

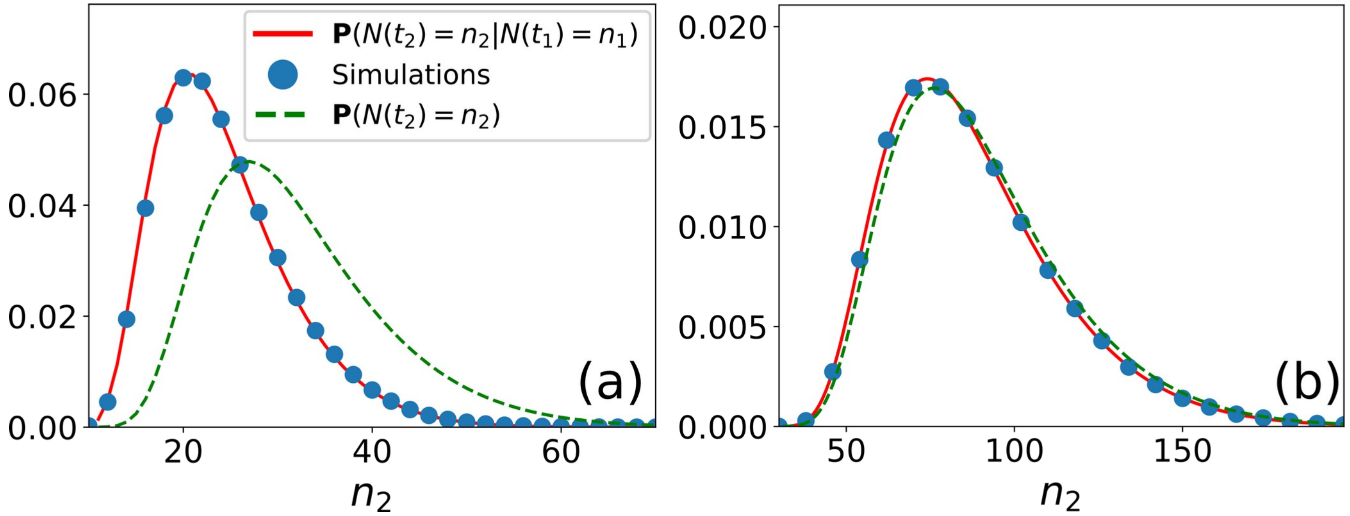


FIG: Conditional distribution for values t_1=200 and n_1=10 and (a) t_2=400 (b) t_2=3200

Is the process Markovian?

Is the process Markovian?

It is not:

$$\mathbb{P}(N(t_3) = n_3 | N(t_2) = n_2, N(t_1) = n_1) \neq \mathbb{P}(N(t_3) = n_3 | N(t_2) = n_2)$$

Is the process Markovian?

It is not:

$$\mathbb{P}(N(t_3) = n_3 | N(t_2) = n_2, N(t_1) = n_1) \neq \mathbb{P}(N(t_3) = n_3 | N(t_2) = n_2)$$

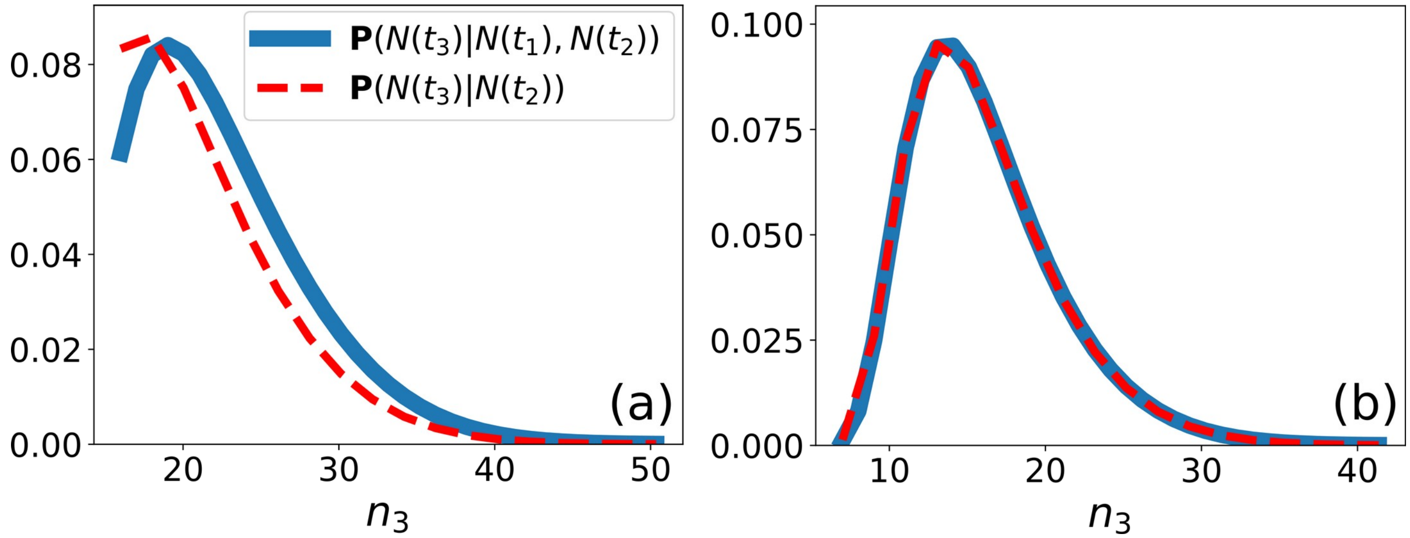


FIG: Distribution of N (t_3 = 200) conditioned on N (t_1 = 100) = 5 and (a) N (t_2 = 110) = 15 and (b) N (t_2 = 110) = 6 (blue curves); and distribution of N (t_3) conditioned only on N (t_2) (red).

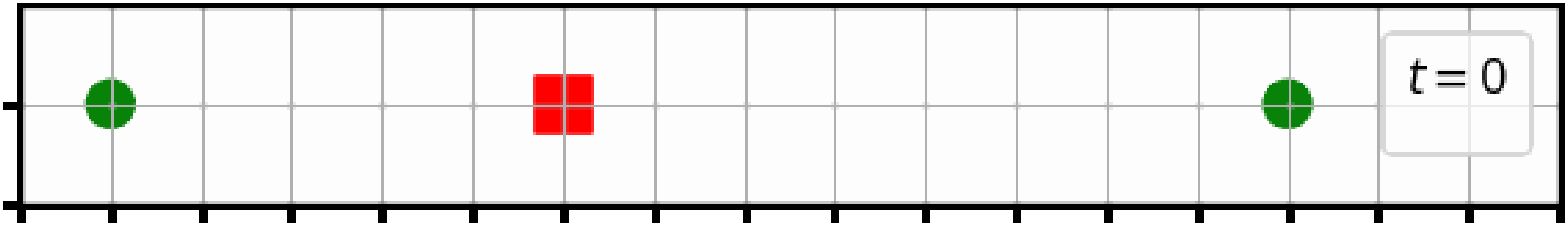
Application to a paradigmatic model: the **trapping problem**.

Application to a paradigmatic model: the **trapping problem**.

Definition: a random walker wanders in an environment which contain at every site a trap with probability c. When the random walk hits a trap, it dies. **(Rosenstock, 1961)**

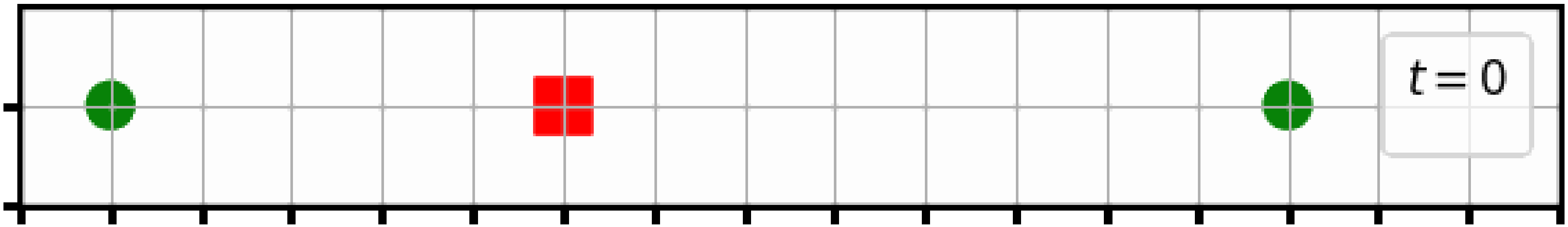
Application to a paradigmatic model: the **trapping problem**.

Definition: a random walker wanders in an environment which contain at every site a trap with probability c. When the random walk hits a trap, it dies. **(Rosenstock, 1961)**



Application to a paradigmatic model: the **trapping problem**.

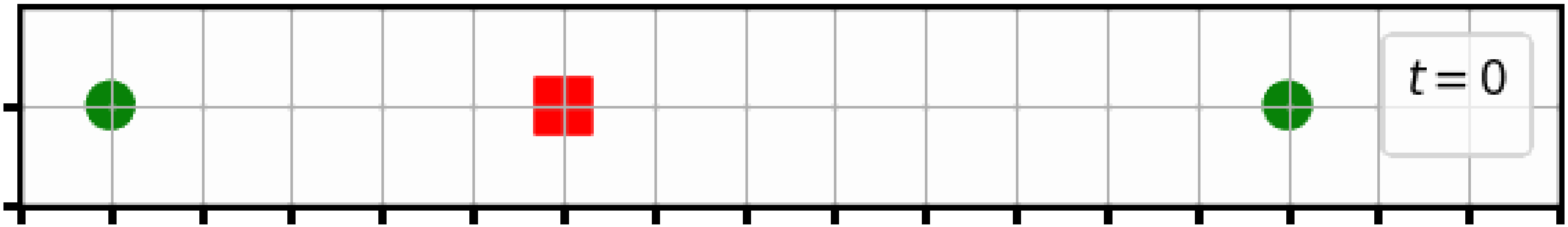
Definition: a random walker wanders in an environment which contain at every site a trap with probability c. When the random walk hits a trap, it dies. **(Rosenstock, 1961)**



For how long does the walker survives?

Application to a paradigmatic model: the **trapping problem**.

Definition: a random walker wanders in an environment which contain at every site a trap with probability c. When the random walk hits a trap, it dies. **(Rosenstock, 1961)**

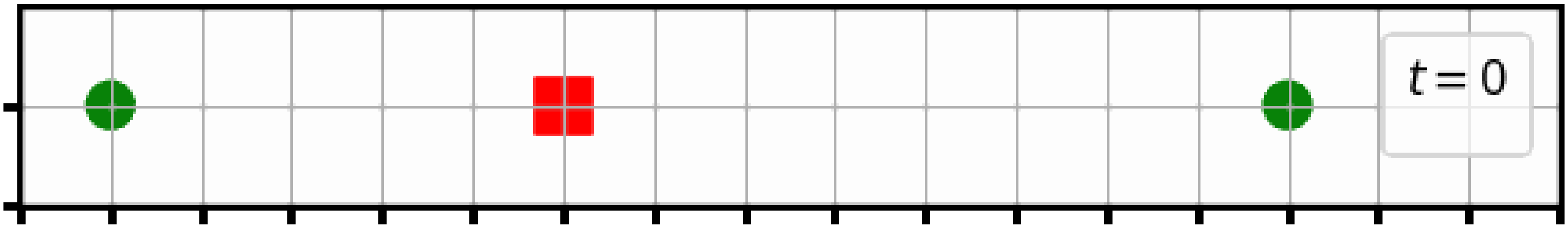


For how long does the walker survives knowing it survived up to t_1?

No known results. Need of the two time quatities: the conditional survival probability is given by the probability to discover no new traps in the newly visited territory.

Application to a paradigmatic model: the **trapping problem**.

Definition: a random walker wanders in an environment which contain at every site a trap with probability c. When the random walk hits a trap, it dies. **(Rosenstock, 1961)**



For how long does the walker survives knowing it survived up to t_1?

No known results. Need of the two time quatities: the conditional survival probability is given by the probability to discover no new traps in the newly visited territory.

$$S(t_2|t_1) = \langle (1-c)^{N(t_2)-N(t_1)} \rangle$$

Surprising feature: there is a non trivial limit at large times, independent of the trap concentration ($N(t)$ exploration process of the 1d brownian motion of variance t)!

Surprising feature: there is a non trivial limit at large times, independent of the trap concentration (N(t) exploration process of the 1d brownian motion of variance t)!

$$\lim_{t_1 \to \infty} S(zt_1|t_1) = \mathbb{P}(N(z) = N(1))$$

Surprising feature: there is a non trivial limit at large times, independent of the trap concentration ($N(t)$ exploration process of the 1d brownian motion of variance t)!

$$\lim_{t_1 \to \infty} S(zt_1|t_1) = \mathbb{P}(N(z) = N(1))$$

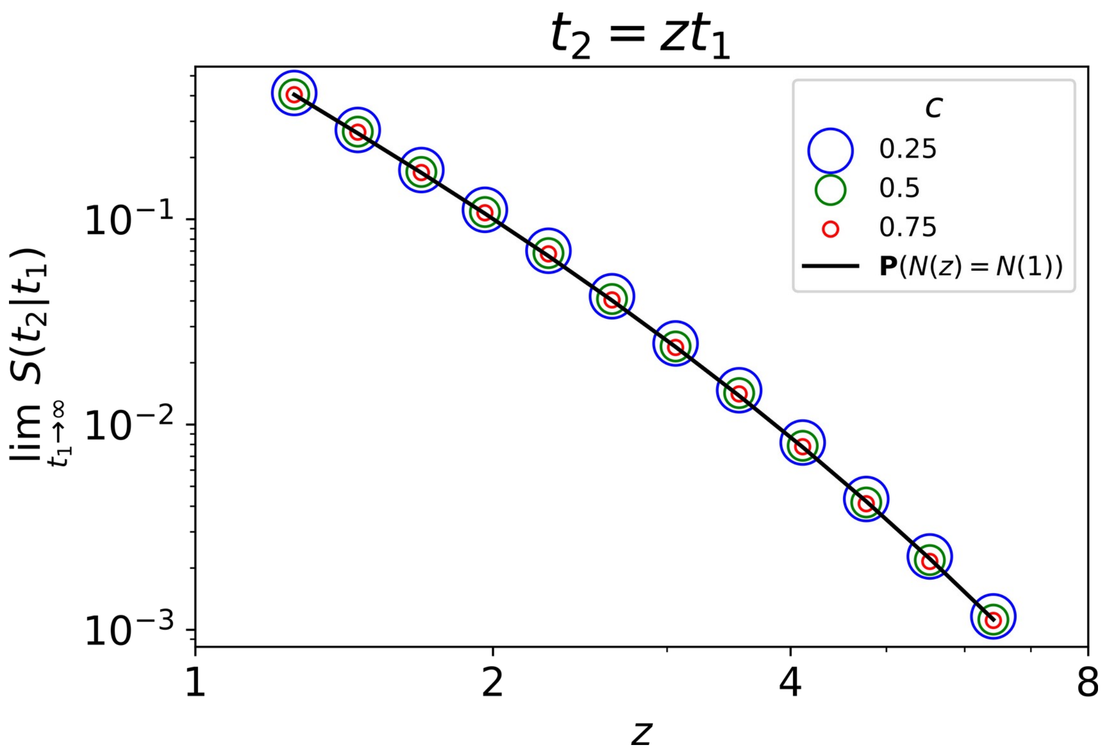


FIG: Conditional survival probability at large times in the limit $t_2=z\, t_1$, z fixed.

For the biased random walks, what about the multiple time distribution?

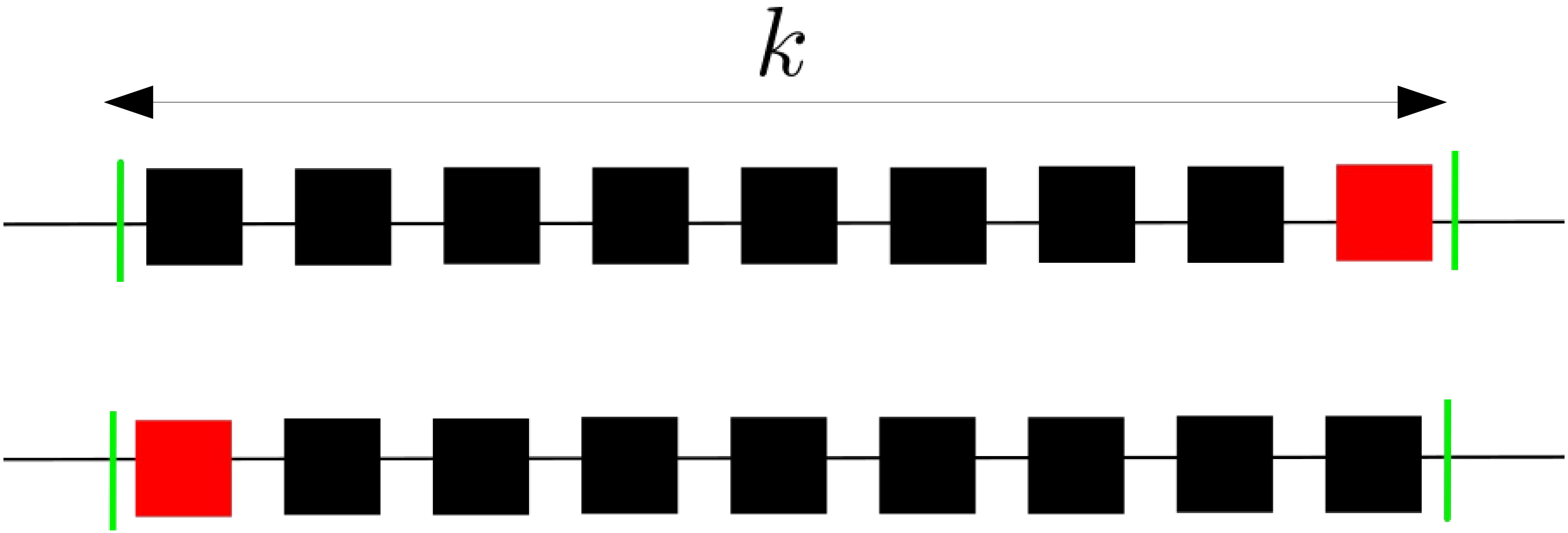
$$\mathbb{P}(N(t_1) = n_1, \ldots, N(t_k) = n_k)$$

We adapt the general steps to obtain the **cumulative two times distribution**, for a symmetric walk to the biased random walk,

$$\mathbb{P}(N(t_1) \geq n_1; N(t_2) \geq n_2)$$

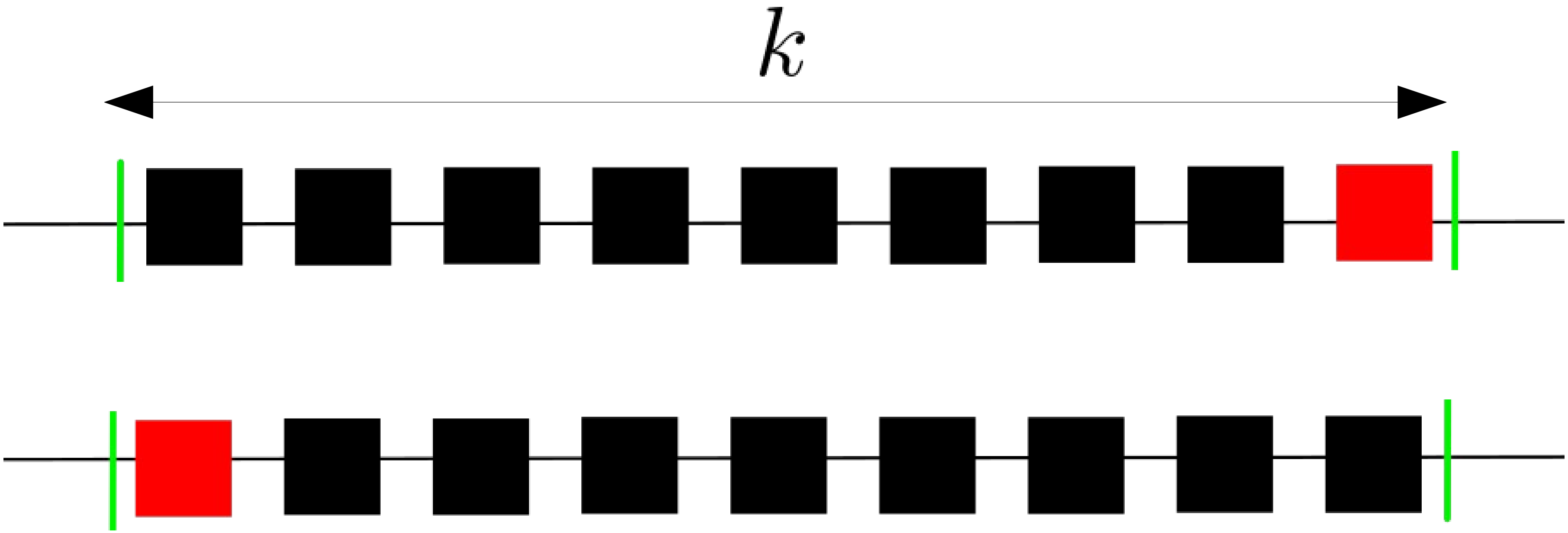
Step 1: We start by solving the exit time problem from an interval, but now the position of start and matters.

Step 1: We start by solving the exit time problem from an interval, but now the position of start and matters.



Step 1: We start by solving the exit time problem from an interval, but now the position of start and matters.

$$\widehat{F}(s,k) \equiv \begin{pmatrix} \widehat{F}^{l\to l}(s,k) & \widehat{F}^{l\to r}(s,k) \\ \widehat{F}^{r\to l}(s,k) & \widehat{F}^{r\to r}(s,k) \end{pmatrix}$$



Step 1: We start by solving the exit time problem from an interval, but now the position matters.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k's, but scalars are replaced by matrices.

$$\mathbb{P}\big(N(t_1) \geq n_1; N(t_2) \geq n_2\big) = \mathbb{P}\big(\tau_0 + ... + \tau_{n_1-1} \leqslant t_1; \tau_0 + ... + \tau_{n_2-1} \leqslant t_2\big)$$

$$\mathcal{L}\{\mathbb{P}(N(t_1) \geq n_1, N(t_2) \geq n_2)\} = \frac{1}{s_1 s_2}\widehat{F}(s_1+s_2,0)\dots\widehat{F}(s_1+s_2,n_1-1)\widehat{F}(s_2,n_1)\dots\widehat{F}(s_2,n_2-1)$$

Step 1: We start by solving the exit time problem from an interval, but now the position matters.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k's, but scalars are replaced by matrices.

Step 3: We take the continuum limit.

We obtain in this limit $\widehat{F}(s,k) = 1 + g(s,k) \quad M(s,m,n) =: \exp\int_m^n g(s,k)dk :$

Step 1: We start by solving the exit time problem from an interval, but now the position matters.

Step 2: We use the key equation, and Laplace transform it in both time variables using the independence of the τ_k's, but scalars are replaced by matrices.

Step 3: We take the continuum limit.

The generalization to the **k times distribution** is now straightforward, following the three steps, valid for biased jumps,

$$\mathcal{L}\left\{\mathbb{P}(N(t_1) \geq n_1; \ldots; N(t_k) \geq n_k)\right\}$$

$$= \frac{1}{2s_1 s_2 \ldots s_k}\,(1,1) \times M(s_1 + \ldots + s_k, 0, n_1) M(s_2 + \ldots + s_k, n_1, n_2) \times \ldots \times M(s_k, n_{k-1}, n_k) \times (1,1)^T\,,$$

... and it can be used to obtain numerical values!

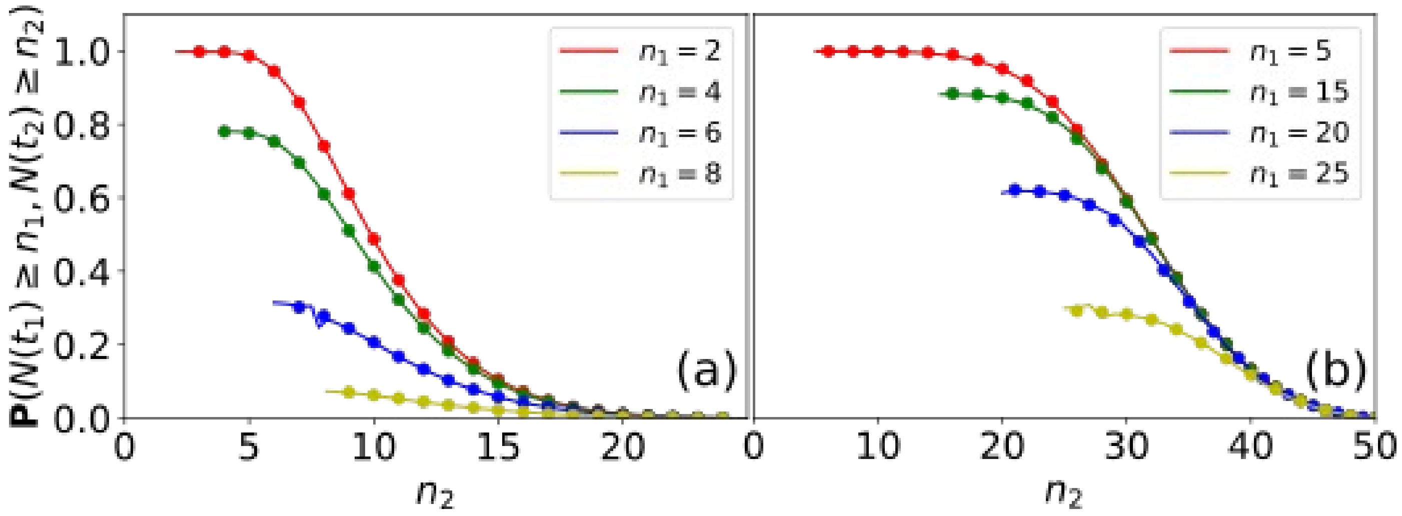


FIG : Numerical inversion of the two time distribution of the number of distinct sites visited for (a) persistent random walk (b) biased random walk

Check the link if you want more details!

The end ...?

- mensions? More general Rws?
- plication to other observables: Perimeter? Holes?

The end ...?

- mensions? More general Rws?
- plication to other observables: Perimeter? Holes?

Check the other link:

The end ...?

- mensions? More general Rws?
- plication to other observables: Perimeter? Holes?
- walks?
- s?

Check the other link:

The end ...?

- mensions? More general Rws?

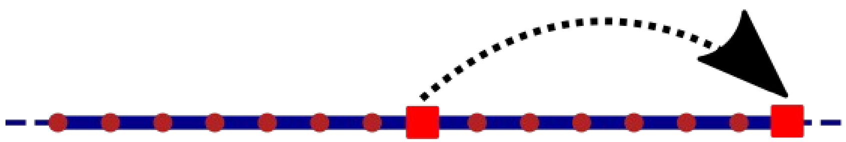
Check the other link:

- plication to other observables: Perimeter? Holes?
- walks?
- s?

Any Questions?

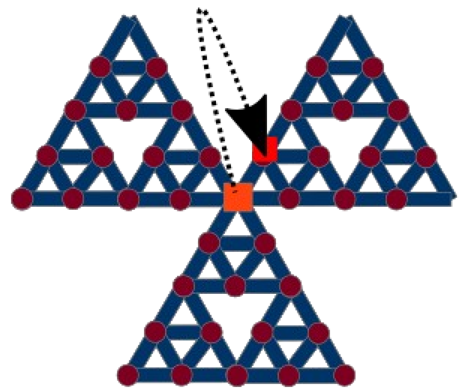
Generalisation:

1d Lévy flight

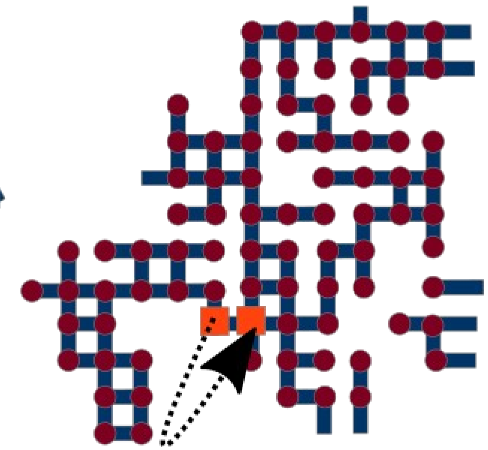


$$p(l) \propto l^{-1-\alpha}$$

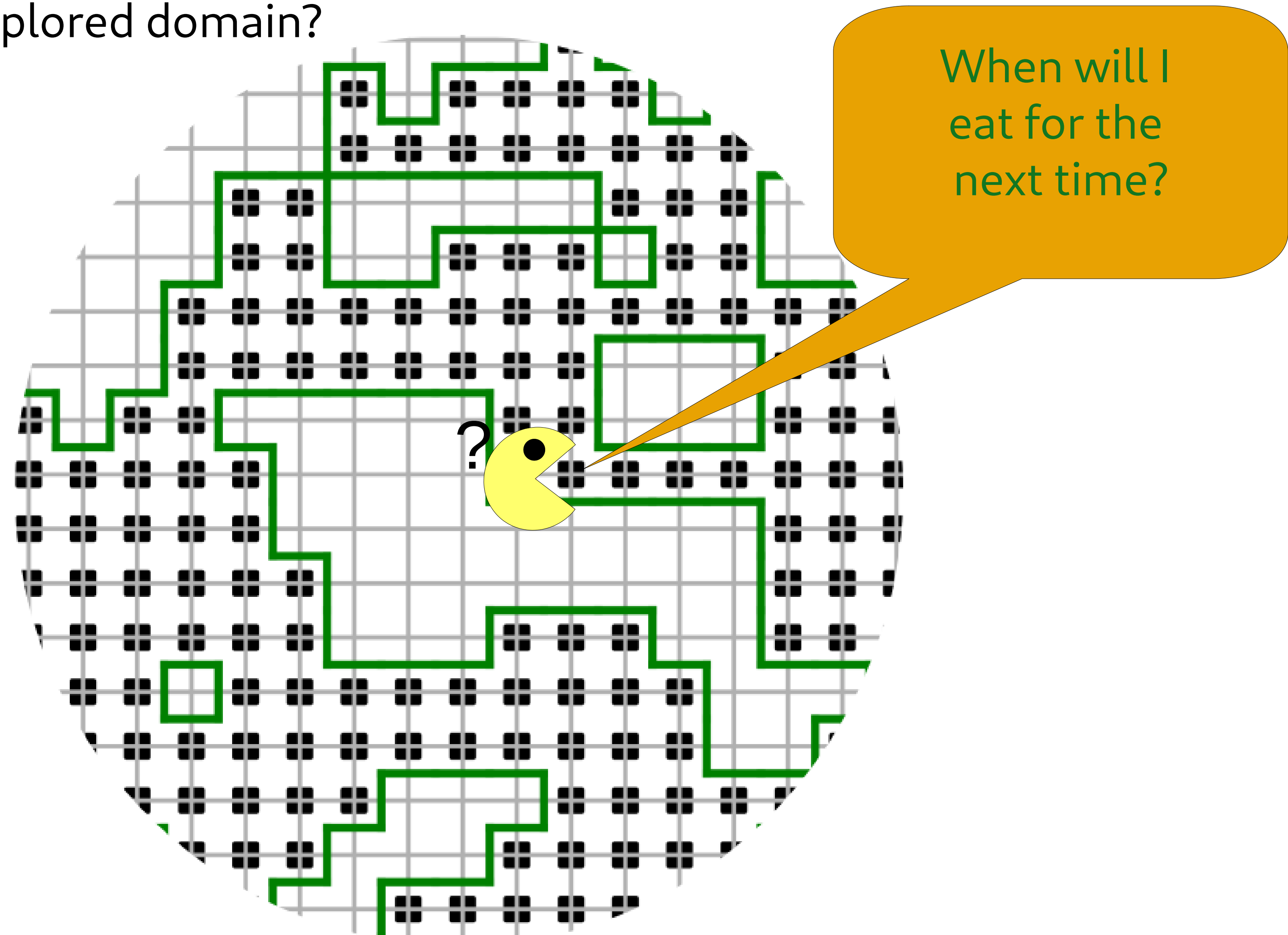
Sierpinski lattice



Percolation cluster



Question: Can we obtain $\mathcal{T}_n$ statistics in the case of more complex geometry of the explored domain?



[1] G. H. Vineyard, J. Math. Phys. 4, 1191 (1963).

[2] K. J. Wiese, J. Stat. Phys. 178, 625 (2020).

[3] B. Annesi, E. Marinari, and G. Oshanin, J. Phys. A: Math. Theor. 52, 345003 (2019).

[4] E. W. Montroll and G. H. Weiss J. of Mathematical Phys. 10, 753 (1969).

[5] R. Rammal, G. Toulouse J. Phys. -Letters 44 (1983).

[6] A. Dembo, The Annals of Probability, Vol. 35, No. 2 , pp. 577-601 (2007).

[7] P. Grassberger and I. Procaccia, J. Chem. Phys. 77, 6281 (1982).