Two-dimensional conditioned simple random walk and some of its surprising properties

Serguei Popov

(based on joint papers with Francis Comets, Nina Gantert, Leonardo Rolla, Daniel Ungaretti, Marina Vachkovskaia)

We write $x \sim y$ if x and y are neighbours in $\mathbb{Z}^d$.

Simple random walk (SRW):

$$P_{xy} = \begin{cases} (2d)^{-1}, & \text{if } x \sim y, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

A fundamental result about SRWs on integer lattices is Pólya's classical theorem:

Theorem

Simple random walk in dimension d is recurrent for $d = 1, 2$ and transient for $d \geq 3$.

A well-known interpretation of this fact, attributed to Shizuo Kakutani, is: "a drunken man always returns home, but a drunken bird will eventually be lost". This observation may explain why birds do not drink vodka.

Still, recurrence in $d = 2$ is *critical*: the return time to the origin is *very* heavy-tailed.

Examples: 1m step size. What are the probabilities of

- going out of Paris
- going out of our galaxy (walking on the *galactic plane*)

before returning to the origin?

See `www.fc.up.pt/pessoas/serguei.popov/2srw.pdf`

From now on, let $(S_n, n \geq 0)$ be two-dimensional simple random walk, transition probabilities given by

$$P_{xy} = \begin{cases} \frac{1}{4}, & \text{if } x \sim y, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Let

$$\tau_0(A) = \inf\{k \geq 0 : S_k \in A\}, \tag{3}$$

$$\tau_1(A) = \inf\{k \geq 1 : S_k \in A\} \tag{4}$$

be the entrance and the hitting time of the set A by simple random walk S (we use the convention $\inf \emptyset = +\infty$). For a singleton $A = \{x\}$, we will write $\tau_i(A) = \tau_i(x)$, $i = 0, 1$, for short.

Potential kernel a of the SRW is defined by

$$a(x) = \sum_{k=0}^{\infty} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big). \tag{5}$$

We have $a(0) = 0$, $a(x) > 0$ for $x \neq 0$.

Also, $\frac{1}{4}\sum_{x\sim 0} a(x) = 1$, which implies by symmetry that

$$a(x) = 1 \text{ for all } x \sim 0. \tag{6}$$

The function a is harmonic outside the origin, i.e.,

$$\frac{1}{4}\sum_{y:y\sim x} a(y) = a(x) \quad \text{for all } x \neq 0, \tag{7}$$

it implies that $a(S_{k\wedge\tau_0(0)})$ is a martingale.

Further, as $x \to \infty$,

$$a(x) = \frac{2}{\pi}\ln\|x\| + \frac{2\gamma + 3\ln 2}{\pi} + O(\|x\|^{-2}) \tag{8}$$

with $\gamma = 0.5772156\ldots$ the Euler-Mascheroni constant.

Define another random walk $(\widehat{S}_n, n \geq 0)$ on $\mathbb{Z}^2 \setminus \{0\}$:

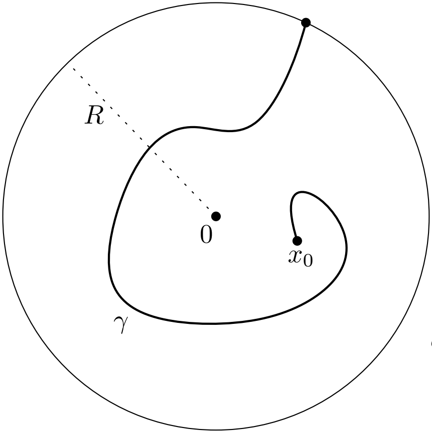
$$\widehat{P}_{xy} = \begin{cases} \dfrac{a(y)}{4a(x)}, & \text{if } x \sim y, x \neq 0, \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

(these are transition probabilities due to (7)).

Let

$$\begin{aligned} \widehat{\tau}_0(A) &= \inf\{k \geq 0 : \widehat{S}_k \in A\} \\ \widehat{\tau}_1(A) &= \inf\{k \geq 1 : \widehat{S}_k \in A\}. \end{aligned}$$

The walk $\widehat{S}$ is the Doob h-transform of the simple random walk, under the condition of not hitting the origin.



$$P_\gamma \approx (1/4)^{|\gamma|} \frac{a(R)}{a(x_0)}$$

$$\mathbb{P}[\text{escape } 0] = q$$

$$a(x_0) \approx q a(R) \quad \text{by O.S.T.}$$

Properties of the walk $\widehat{S}$:

(i) $\widehat{S}$ is reversible, with the reversible measure $\mu_x := a^2(x)$. on the two-dimensional lattice with conductances $\big(a(x)a(y), x, y \in \mathbb{Z}^2, x \sim y\big)$. neighbours of the origin. Then the process $1/a(\widehat{S}_{k\wedge\widehat{\tau}_0(\mathcal{N})})$ is a martingale. $\widehat{S}$ is transient. Moreover, for all $x \neq 0$, $x \neq y$, $x, y \neq 0$

$$\mathbb{P}_x\big[\widehat{\tau}_1(x) < \infty\big] = 1 - \frac{1}{2a(x)}, \text{ and}$$

$$\mathbb{P}_x\big[\widehat{\tau}_0(y) < \infty\big] = \mathbb{P}_x\big[\widehat{\tau}_1(y) < \infty\big] = \frac{a(x) + a(y) - a(x-y)}{2a(x)}.$$

It is curious to observe that

$$\mathbb{P}_x\big[\widehat{\tau}_1(y) < \infty\big] = \frac{a(x) + a(y) - a(x-y)}{2a(x)}$$

implies that, for any x, $\mathbb{P}_x[\widehat{\tau}_1(y) < \infty]$ converges to $\frac{1}{2}$ as $y \to \infty$. (Since $\widehat{S}$ is transient and there are "many ways to escape" on the plane, one could naturally think that the above limit would be 0.)

Notations: define the (discrete) ball

$$\mathsf{B}(x,r) = \{y \in \mathbb{Z}^2 : \|y - x\| \le r\}$$

and we abbreviate $\mathsf{B}(r) := \mathsf{B}(0,r)$.
The (internal) boundary of $A \subset \mathbb{Z}^2$ is defined by

$$\partial A = \{x \in A : \text{there exists } y \in \mathbb{Z}^2 \setminus A \text{ such that } x \sim y\}.$$

For a set $T \subset \mathbb{Z}_+$ (thought of as a set of time moments), let

$$\widehat{S}_T = \bigcup_{m \in T} \{\widehat{S}_m\}$$

be the *range* of the walk $\widehat{S}$ with respect to that set.

For a nonempty and finite set $A \subset \mathbb{Z}^2$, let us consider random variables

$$\mathcal{R}(A) = \frac{\big|A \cap \widehat{S}_{[0,\infty)}\big|}{|A|},$$
$$\mathcal{V}(A) = \frac{\big|A \setminus \widehat{S}_{[0,\infty)}\big|}{|A|} = 1 - \mathcal{R}(A);$$

that is, $\mathcal{R}(A)$ (respectively, $\mathcal{V}(A)$) is the proportion of visited (respectively, unvisited) sites of A by the walk $\widehat{S}$ (assume it starts at the origin).
Abbreviate, for $M_0 > 0$,

$$\ell_A = |A|^{-1} \max_{y \in A} \big|A \cap \mathrm{B}\big(y, \tfrac{n}{\ln^{M_0} n}\big)\big|. \tag{10}$$

Theorem

Let $M_0 > 0$ be a fixed constant, and assume that $A \subset \mathrm{B}(n) \setminus \mathrm{B}(n \ln^{-M_0} n)$. Then, for all $s \in [0,1]$, we have, with positive constants $c_{1,2}$ depending only on M_0,

$$\left|\mathbb{P}[\mathcal{V}(A) \leq s] - s\right| \leq c_1 \left(\frac{\ln \ln n}{\ln n}\right)^{1/3} + c_2 \ell_A \left(\frac{\ln \ln n}{\ln n}\right)^{-2/3}, \quad (11)$$

and the same result holds with $\mathcal{R}$ on the place of $\mathcal{V}$.

The above result means that if $A \subset \mathrm{B}(n) \setminus \mathrm{B}(\varepsilon_0 n)$ is “big enough and well distributed”, then the proportion of visited sites has approximately Uniform$[0,1]$ distribution.
In particular, on can obtain the following

Corollary

Assume that $D \subset \mathbb{R}^2$ is a bounded open set. Then both sequences $(\mathcal{R}(nD \cap \mathbb{Z}^2), n \geq 1)$ and $(\mathcal{V}(nD \cap \mathbb{Z}^2), n \geq 1)$ converge in distribution to the Uniform$[0,1]$ random variable.

Indeed, it is straightforward to obtain it from the theorem since (note that D contains a disk) $|nD \cap \mathbb{Z}^2|$ is of order n^2 as $n \to \infty$, and so $\ell_{nD\cap\mathbb{Z}^2}$ will be of order $\ln^{-2M_0} n$, with an arbitrarily large M_0.

Also, we prove that the range of $\widehat{S}$ contains many "big holes". To formulate this result, we need the following

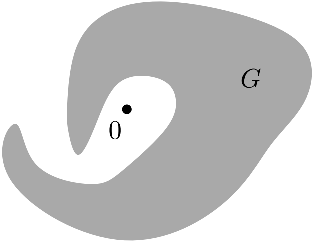
Definition

We say that a set $G \subset \mathbb{R}^2$ does not surround the origin, if

- *there exists $c_1 > 0$ such that $G \subset \mathrm{B}(c_1)$, i.e., G is bounded;*
- *there exist $c_{2,3} > 0$ and a function $f = (f_1, f_2) : [0,1] \mapsto \mathbb{R}^2$ such that $f(0) = 0$, $\|f(1)\| = c_1$, $|f_j'(s)| \leq c_2$ for all $s \in [0,1]$, and*

$$\inf_{s \in [0,1], y \in G} \|(f_1(s), f_2(s)) - y\| \geq c_3,$$

i.e., one can escape from the origin to infinity along a path which is uniformly away from G.



Theorem

Let $G \subset \mathbb{R}^2$ be a set that does not surround the origin. Then,

$$\mathbb{P}\left[nG \cap \widehat{S}_{[0,\infty)} = \emptyset \text{ for infinitely many } n\right] = 1. \tag{12}$$

Recall that a set is called *recurrent* with respect to the Markov chain, if it is visited infinitely many times almost surely; a set is called *transient*, if it is visited only finitely many times almost surely.

Note that, in general, a set can be neither recurrent nor transient — think e.g. of the simple random walk on a binary tree, fix a neighbor of the root and consider the set of vertices of the tree connected to the root through this fixed neighbor.

It is clear that any nonempty set is recurrent with respect to a recurrent Markov chain, and every finite set is transient with respect to a transient Markov chain.

In many situations it is possible to characterize completely the recurrent and transient sets, as well as to answer the question if any set must be either recurrent or transient.

For example, for the simple random walk in $\mathbb{Z}^d$, $d \geq 3$, each set is either recurrent or transient and the characterization is provided by *Wiener's test*, formulated in terms of capacities of intersections of the set with exponentially growing annuli.

Now, for the conditioned two-dimensional walk $\widehat{S}$ the characterization of recurrent and transient sets is particularly simple:

Theorem

A set $A \subset \mathbb{Z}^2$ is recurrent with respect to $\widehat{S}$ if and only if A is infinite.

Let us define the "future minimum"

$$M_n := \min_{m \geq n} \|\widehat{S}_m\|;$$

by transience, $M_n \to \infty$ a.s.

Theorem

For every $0 < \delta < \frac{1}{2}$ we have, almost surely,

$$M_n \leq n^\delta \text{ i.o.} \quad \text{but} \quad M_n \geq \frac{\sqrt{n}}{\ln^\delta n} \text{ i.o.}$$

and, on the other hand,

$$e^{\ln^{1-\delta} n} \leq M_n \leq \sqrt{n} \times \sqrt{(e+\delta)(\ln \ln n)} \quad \text{eventually.}$$

Encounters: consider two independent $\widehat{S}$-walks, denoted $\widehat{S}^1$ and $\widehat{S}^2$, and also an independent SRW S.

Theorem

Let $x_1, x_2 \in \mathbb{Z}^2 \setminus \{0\}$ have the same parity. Then, we have

$$\mathbb{P}_{x_1,x_2}\big[\widehat{S}^1_n = \widehat{S}^2_n \text{ i.o.}\big] = 1,$$

and

$$\mathbb{P}_{x_1,x_2}\big[\widehat{S}^1_n = S_n \text{ i.o.}\big] = 1.$$

We abbreviate $\tau_1(R) = \tau_1(\partial B(R))$. We will consider, with a slight abuse of notation, the function

$$a(r) = \frac{2}{\pi}\ln r + \frac{2\gamma + 3\ln 2}{\pi}$$

of a *real* argument $r \geq 1$. Then, we get

$$\sum_{y \in \partial \mathsf{B}(x,r)} \nu(y) a(y) = a(r) + O\Big(\frac{\|x\| \vee 1}{r}\Big) \tag{13}$$

for *any* probability measure ν on $\partial \mathsf{B}(x,r)$.

Using optional stopping theorem applied to the martingale $a(S_{n\wedge\tau_0(0)})$, we obtain for all $x \in \mathbb{Z}^2$ and $R \geq 1$ such that $x, y \in \mathrm{B}(R)$, $x \neq y$

$$\mathbb{P}_x[\tau_1(R) < \tau_1(y)] = \frac{a(x-y)}{a(R) + O\big(R^{-1}(\|y\| \vee 1)\big)}, \qquad (14)$$

as $R \to \infty$.

Also, using martingale $1/a(\widehat{S}_{k\wedge\widehat{\tau}_0(\mathcal{N})})$ yields

$$\mathbb{P}_x[\widehat{\tau}_1(R) < \widehat{\tau}_1(r)] = \frac{(a(r))^{-1} - (a(x))^{-1} + O(R^{-1})}{(a(r))^{-1} - (a(R))^{-1} + O(r^{-1})}, \qquad (15)$$

for $1 < r < \|x\| < R < \infty$.
Sending R to infinity in (15) we see that for $1 \leq r \leq \|x\|$

$$\mathbb{P}_x[\widehat{\tau}_1(r) = \infty] = 1 - \frac{a(r) + O(r^{-1})}{a(x)}. \qquad (16)$$

We also need (and have) the fact that the walks S and $\widehat{S}$ are almost indistinguishable on a “distant” (from the origin) set.

Refined bounds on the hitting probabilities for excursions of the conditioned walk:

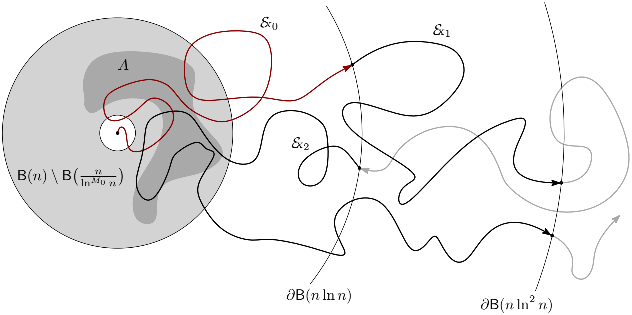


Figure: Excursions and their visits to A

Let us assume that $\|x\| \geq n \ln^{-M_0} n$ and $y \in A$, where the set A is as in Theorem 3.1. Also, abbreviate $R = n \ln^2 n$.

Lemma

In the above situation, we have

$$\begin{aligned} \mathbb{P}_x[\widehat{\tau}_1(y) < \widehat{\tau}_1(R)] &= \big(1 + O(\ln^{-3} n)\big) \\ &\quad \times \frac{a(x)a(R) + a(y)a(R) - a(x-y)a(R) - a(x)a(y)}{a(x)(2a(R) - a(y))}. \end{aligned} \tag{17}$$

Theorem

Let $M_0 > 0$ be a fixed constant, and assume that $A \subset \mathrm{B}(n) \setminus \mathrm{B}(n \ln^{-M_0} n)$. Then, for all $s \in [0,1]$, we have, with positive constants $c_{1,2}$ depending only on M_0,

$$\big|\mathbb{P}[\mathcal{V}(A) \leq s] - s\big| \leq c_1 \Big(\frac{\ln \ln n}{\ln n}\Big)^{1/3} + c_2 \ell_A \Big(\frac{\ln \ln n}{\ln n}\Big)^{-2/3}, \quad (18)$$

and the same result holds with $\mathcal{R}$ on the place of $\mathcal{V}$.

Idea:
We consider the visits to the set A during excursions of the walk from $\partial B(n \ln n)$ to $\partial B(n \ln^2 n)$, see Figure 1.

The crucial argument is the following: the randomness of $\mathcal{V}(A)$ comes from the *number* of excursions and not from the excursions *themselves*. If the number of excursions is around $c \times \frac{\ln n}{\ln \ln n}$, then it is possible to show (using a standard weak-LLN argument) that the proportion of uncovered sites in A is *concentrated* around e^{-c}. On the other hand, that number of excursions can be modeled roughly as $Y \times \frac{\ln n}{\ln \ln n}$, where Y is an Exponential(1) random variable.

Then, $\mathbb{P}[\mathcal{V}(A) \le s] \approx \mathbb{P}[Y \ge \ln s^{-1}] = s$, as required.

Theorem

A set $A \subset \mathbb{Z}^2$ is recurrent with respect to $\widehat{S}$ if and only if A is infinite.

We only need to prove that every infinite subset of $\mathbb{Z}^d$ is recurrent for $\widehat{S}$. Basically, this is a consequence of the fact that, due to

$$\mathbb{P}_x\big[\widehat{\tau}_0(y) < \infty\big] = \mathbb{P}_x\big[\widehat{\tau}_1(y) < \infty\big] = \frac{a(x) + a(y) - a(x-y)}{2a(x)}.$$

we have

$$\lim_{y\to\infty} \mathbb{P}_{x_0}\big[\widehat{\tau}_1(y) < \infty\big] = \frac{1}{2} \tag{19}$$

for any $x_0 \in \mathbb{Z}^2$.

Indeed, let $\widehat{S}_0 = x_0$; since A is infinite, by (19) one can find $y_0 \in A$ and R_0 such that $\{x_0, y_0\} \subset \mathrm{B}(R_0)$ and

$$\mathbb{P}_{x_0}\big[\widehat{\tau}_1(y_0) < \widehat{\tau}_1(R_0)\big] \geq \frac{1}{3}.$$

Then, for any $x_1 \in \partial\mathrm{B}(R_0)$, we can find $y_1 \in A$ and $R_1 > R_0$ such that $y_1 \in \mathrm{B}(R_1) \setminus \mathrm{B}(R_0)$ and

$$\mathbb{P}_{x_1}\big[\widehat{\tau}_1(y_1) < \widehat{\tau}_1(R_1)\big] \geq \frac{1}{3}.$$

Continuing in this way, we can construct a sequence $R_0 < R_1 < R_2 < \dots$ (depending on the set A) such that, for each $k \geq 0$, the walk $\widehat{S}$ hits A on its way from $\partial \mathrm{B}(R_k)$ to $\partial \mathrm{B}(R_{k+1})$ with probability at least $\frac{1}{3}$, regardless of the past. This clearly implies that A is a recurrent set.

Main references:

F. COMETS, S. POPOV, M. VACHKOVSKAIA (2016) Two-dimensional random interlacements and late points for random walks. *Commun. Math. Phys.* **343**, 129–164.

N. GANTERT, S. POPOV, M. VACHKOVSKAIA (2019) On the range of a two-dimensional conditioned simple random walk. *Ann. H. Lebesgue* **2**, 349–368.

S. POPOV, L.T. ROLLA, D. UNGARETTI (2020) Transience of conditioned walks on the plane: encounters and speed of escape. *Electron. J. Probab.* **25**, paper no. 52.

S. POPOV (2021) *Two-Dimensional Random Walk: From Path Counting to Random Interlacements*. Cambridge University Press, Cambridge.

Questions?