On random walks on the random graph

Laurent Miclo
(Toulouse School of Economics)

Joint work with Sourav Chatterjee and Persi Diaconis
(Stanford University)

Plan of the talk

1. Sphere random walks
2. On the random graph
3. Spectral gap
4. A Lyapounov argument
5. Hardy's bounds
6. On a deterministic Rado graph

Plan

The random graph

Consider the random graph $\mathcal{G} := (\mathbb{Z}_+, E)$, where each undirected edge $\{x, y\} \subset \mathbb{Z}_+$, with $x \neq y$, belongs to E with probability $1/2$, independently for all of them.

When the probability $1/2$ is replaced by $\rho \in (0, 1)$, the corresponding notions will receive ρ in index.

For any $x \in \mathbb{Z}_+$, define

$$N(x) \quad := \quad \{y \in \mathbb{Z}_+ \ : \ \{x, y\} \in E\}$$

the sphere of radius 1 around x for $\mathcal{G}$.

An associated Markov kernel

Consider the probability measure Q given on $\mathbb{Z}_+$ by

$$\forall\ x \in \mathbb{Z}_+, \qquad Q(x) \;:=\; \frac{1}{2^{1+x}}$$

and associate a Markov kernel K on $\mathbb{Z}_+$ via

$$\forall\ x, y \in \mathbb{Z}_+, \qquad K(x,y) \;:=\; \frac{Q(y)}{Q(N(x))}\mathbb{1}_{N(x)}(y)$$

This kernel is reversible to the probability π given by

$$\forall\ x \in \mathbb{Z}_+, \qquad \pi(x) \;=\; Z^{-1}Q(x)Q(N(x))$$

where $Z > 0$ is the normalizing (random) constant.

Sphere "random walks"

For $x \in \mathbb{Z}_+$, let $(X_t^x)_{t \geqslant 0}$ be a Markov process starting from x and whose generator is $K - \mathrm{Id}$.
We are interested in its speed of convergence to π. Define the mixing time

$$\tau^x \quad := \quad \min\left\{t \geqslant 0 \,:\, \|\mathcal{L}(X_t^x) - \pi\|_{\mathrm{tv}} \leqslant \frac{1}{2}\right\}$$

The main result of the talk is:

Theorem 1

There exist two (random) constants $a, b > 0$ such that for any $x \in \mathbb{Z}_+$,

$$\tau^x \quad \leqslant \quad b(1 + \log_a^*(x))$$

Iterated logarithms

Recall that for $a > 0$, the iterated logarithm $\log_a^*$ is defined as follows.
Consider the smallest $x_a \geqslant 0$ such that for any $x \geqslant x_a$, we have

$$\log_a(x) := \frac{\log(x)}{\log(a)} \leqslant x$$

If $x \in [0, x_a]$, by definition, we take $\log_a^*(x) = 0$. Otherwise $\log_a^*(x)$ is the minimal number of time one has to iterate $\log_a$, starting from x, to get a number below x_a. Namely

$$\underbrace{\log_a \circ \log_a \circ \cdots \circ \log_a}_{\log_a^* \text{ times}}(x) \leqslant x_a$$

The function $\log_a^*$ grows very slowly, as mentioned by Persi.

Plan

The parent vertex

For any $x \in \mathbb{Z}_+$, define the "parent" vertex of x by

$$p(x) \quad := \quad \min N(x) \quad = \quad \arg\max_{N(x)} Q$$

and consider the event

$$\mathcal{B} \quad := \quad \{\forall\ x \in \mathbb{N},\ p(x) < x\}$$

It happens with positive probability:

Lemma 2

We have $\mathbb{P}[\mathcal{B}] \geqslant 1/4$. More generally, for any $\rho \in (0,1)$, we have $\mathbb{P}_\rho[\mathcal{B}] > 0$.

Proof of Lemma 2 (1)

For the second statement, define A_x, for $x \in \mathbb{N}$, as the event that x is not linked in $\mathcal{G}_\rho$ to a smaller vertex. Namely,

$$A_x := \bigcap_{y \in [\![0, x-1]\!]} \{B_{\{y,x\}} = 0\}$$

These events are independent and $\mathbb{P}[A_x] = (1-\rho)^x$. We have

$$\mathcal{B} = \bigcap_{x \in \mathbb{N}} A_x^{\mathrm{c}}$$

Simple computations lead to

$$\mathbb{P}_\rho[\mathcal{B}] = \left(\sum_{n \in \mathbb{Z}_+} p(n)(1-\rho)^n \right)^{-1}$$

where $p(n)$ is the number of partitions of n. Since this quantity behaves like an exponential of $\sqrt{n}$ for large n, we get $\mathbb{P}_\rho[\mathcal{B}] > 0$.

Proof of Lemma 2 (2)

For the first bound, we could try to use upper bound on the partition numbers. It is simpler to use Kounias-Hunter-Worsley bound for unions of pairwise independent events:

$$\mathbb{P}\left[\bigcup_{x\in[\![n]\!]} A_x\right] \leqslant 1 \wedge \left(\sum_{x\in[\![n]\!]} \mathbb{P}[A_x] - \mathbb{P}[A_1] \sum_{y\in[\![2,n]\!]} \mathbb{P}[A_y]\right)$$

using that

$$\mathbb{P}[A_1] \geqslant \mathbb{P}[A_2] \geqslant \cdots \geqslant \mathbb{P}[A_n]$$

We get

$$\begin{aligned} \mathbb{P}\left[\mathcal{B}^{c}\right] &= \mathbb{P}\left[\bigcup_{x\in\mathbb{N}} A_x\right] \\ &\leqslant \lim_{n\to\infty} \frac{1}{2} + \frac{1}{4} - \frac{1}{2^{n+1}} = \frac{3}{4} \end{aligned}$$

A tree

Consider the set of edges

$$F \quad := \quad \{\{x, p(x)\} \; : \; x \in \mathbb{N}\}$$

and the corresponding graph $\mathcal{T} := (\mathbb{Z}_+, F)$.
Under $\mathcal{B}$, it is clear that $\mathcal{T}$ is a tree. In fact this is always true:

Lemma 3

The graph $\mathcal{T}$ is a tree.

Proof

To show $\mathcal{T}$ does not contain cycles, note that when $y = p(x)$ and $z = p(y)$, then $z < x$, because $\{x, z\} \subset N(y)$.
Furthermore $\mathcal{T}$ is connected, since following p, one ends up being decreasing and attaining 0.

$\blacksquare$

Lemma 4

A.s. there exists only a finite number of $x \in \mathbb{N}$ such that $p(x) > 2\log_2(1+x)$. In particular, a.s. there exists a finite (random) $K \geqslant 2$ such that

$$\forall\ x \in \mathbb{N}, \qquad p(x) \leqslant K\log(1+x)$$

Proof

$$\begin{aligned}
&\sum_{x\in\mathbb{N}} \mathbb{P}[p(x) > 2\log(1+x)] \\
&\quad = \sum_{x\in\mathbb{N}} \mathbb{P}[B_{\{0,x\}} = 0, B_{\{1,x\}} = 0, \ldots, B_{\{\lfloor 2\log(1+x)\rfloor,x\}} = 0] \\
&\quad = \sum_{x\in\mathbb{N}} \frac{1}{2^{1+\lfloor 2\log(1+x)\rfloor}} \\
&\quad \leqslant \sum_{x\in\mathbb{N}} \frac{1}{(1+x)^2} < +\infty \qquad \blacksquare
\end{aligned}$$

Plan

Spectral gap

Our goal is to obtain a positive spectral gap, first under $\mathcal{B}$.

Proposition 5

On $\mathcal{B}$, there exists a random constant $\Lambda > 0$ such that

$$\forall\ f \in \mathbb{L}^2(\pi), \qquad \Lambda\pi[(f - \pi[f])^2] \;\leqslant\; \mathcal{E}(f)$$

where in the r.h.s. $\mathcal{E}$ is the Dirichlet form defined by

$$\forall\ f \in \mathbb{L}^2(\pi), \qquad \mathcal{E}(f) \;\coloneqq\; \frac{1}{2} \sum_{x,y \in \mathbb{Z}_+} (f(y) - f(x))^2 \pi(x) K(x,y)$$

First Dirichlet eigenvalue

Since $\pi[(f - \pi[f])^2] \leqslant \pi[(f - f(0))^2]$, the previous result is an immediate consequence of the existence of a positive first Dirichlet eigenvalue:

Proposition 6

On $\mathcal{B}$, there exists a random constant $\Lambda > 0$ such that

$$\forall\ f \in \mathbb{L}^2(\pi), \qquad \Lambda\pi[(f - f[0])^2] \ \leqslant\ \mathcal{E}(f) \tag{1}$$

The proof of Proposition 6 is based on the pruning of $\mathcal{G}$ into $\mathcal{T}$ and on the resort to Cheeger's inequalities for trees.

Define the Markov kernel $K_{\mathcal{T}}$ via

$$\forall\ x, y \in \mathbb{Z}_+,\ \ K_{\mathcal{T}}(x,y) \quad := \quad \begin{cases} K(x,y) & \text{, if } \{x,y\} \in F \\ 1 - \sum_{z\in\mathbb{Z}_+\setminus\{x\}} K_{\mathcal{T}}(x,z) & \text{, if } x = y \\ 0 & \text{, otherwise} \end{cases}$$

the Dirichlet form $\mathcal{E}_{\mathcal{T}}$ given by

$$\forall\ f \in \mathbb{L}^2(\pi), \qquad \mathcal{E}_{\mathcal{T}}(f) \quad = \quad Z \sum_{\{x,y\}\in F} (f(y) - f(x))^2\, \pi(x) K(x,y)$$

and the (non-negative) measure μ through

$$\forall\ x \in \mathbb{N}, \qquad \mu(x) \quad := \quad Q(x)Q(p(x)) \tag{2}$$

Pruning (2)

Proposition 7

On $\mathcal{B}$, there exists $\lambda > 0$ such that

$$\forall\ f \in \mathbb{L}^2(\mu), \qquad \lambda\mu[(f - f(0))^2] \ \leqslant\ \mathcal{E}_{\mathcal{T}}(f) \tag{3}$$

Proposition 6 follows with $\Lambda := \lambda/2$, due to the inclusion $N(x) \subset [\![p(x), \infty[\![$ and to the exponential feature of Q implying

$$\forall\ x \in \mathbb{N}, \qquad Q(p(x)) \ \leqslant\ Q(N(x)) \ \leqslant\ 2Q(p(x))$$

Indeed,

$$\begin{aligned}
\lambda\pi[(f - f(0))^2] &= \frac{\lambda}{Z}\sum_{x\in\mathbb{N}}(f(x) - f(0))^2 Q(x)Q(N(x)) \\
&\leqslant \frac{2\lambda}{Z}\sum_{x\in\mathbb{N}}(f(x) - f(0))^2 Q(x)Q(p(x)) \\
&= \frac{2\lambda}{Z}\mu[(f - f(0))^2] \ =\ \frac{2}{Z}\mathcal{E}_{\mathcal{T}}(f) \ \leqslant\ 2\mathcal{E}(f)
\end{aligned}$$

Dirichlet-Cheeger

For any $A \subset \mathbb{N}$, define $\partial A := \{\{x, y\} : x \in A, y \notin A\}$. Endow the set of edges with the measure ν induced by

$$\nu(\{x, y\}) \quad := \quad Z\pi(x)K_{\mathcal{T}}(x, y)$$

Define the Dirichlet-Cheeger constant

$$\iota \quad := \quad \inf_{A \in \mathcal{A}} \frac{\nu(\partial A)}{\mu(A)} \quad \geqslant \quad 0$$

where

$$\mathcal{A} \quad := \quad \{A \subset \mathbb{N} : A \neq \emptyset\}$$

The Dirichlet-Cheeger inequality states

$$\lambda \quad \geqslant \quad \frac{\iota^2}{2}$$

Proposition 8

On $\mathcal{B}$, we have $\iota \geqslant 1/2$.

Proof

Decomposing an element of $\mathcal{A}$ into its $\mathcal{T}$-connected components and including each component into the substree $\mathcal{T}_a$ generated its smallest element a, we get

$$\begin{aligned} \iota &= \inf_{a \in \mathbb{N}} \frac{\nu(\partial \mathcal{T}_a)}{\mu(\mathcal{T}_a)} \\ &= \inf_{a \in \mathbb{N}} \frac{Q(a)Q(p(a))}{\mu(\mathcal{T}_a)} \end{aligned}$$

Note that on $\mathcal{B}$,

$$\begin{array}{rrcl} \forall\ x \in \mathcal{T}_a, & p(x) & \geqslant & p(a) \\ & \mathcal{T}_a & \subset & [\![a, \infty[\![\end{array}$$

We deduce

$$\begin{array}{rcl} \mu(\mathcal{T}_a) & = & \displaystyle\sum_{x \in \mathcal{T}_a} Q(x) Q(p(x)) \\ & \leqslant & Q(p(a)) \displaystyle\sum_{x \in \mathcal{T}_a} Q(x) \\ & \leqslant & Q(p(a)) \displaystyle\sum_{x \in [\![a, \infty[\![} Q(x) \\ & = & 2Q(p(a))Q(a) \end{array}$$

$\blacksquare$

Ouside $\mathcal{B}$

In the general case, note from Lemma 4 that there exists a (random) vertex $x_0 \in \mathbb{Z}_+$ such that

$$\forall\ x > x_0, \qquad p(x) \quad < \quad x$$

We deduce there exists $x_1 \geqslant x_0$ such that

$$\forall\ a > x_1,\ \forall\ x \in \mathcal{T}_a, \qquad p(x) \quad < \quad x$$

and from the above proof

$$\inf_{a > x_1} \frac{\nu(\partial \mathcal{T}_a)}{\mu(\mathcal{T}_a)} \quad \geqslant \quad \frac{1}{2}$$

By finiteness of $[\![x_1]\!]$, we also have

$$\inf_{a \in [\![x_1]\!]} \frac{\nu(\partial \mathcal{T}_a)}{\mu(\mathcal{T}_a)} \quad > \quad 0$$

and thus $\iota > 0$.

Plan

Coming back close to 0

The spectral gap and Lemma 4 imply the bound

$$\begin{aligned}\|\mathcal{L}(X_t^x) - \pi\|_{\mathrm{tv}} &\leqslant \frac{1}{\sqrt{\pi(x)}} e^{-\Lambda t} \\ &\leqslant (x+1)^K 2^{x/2} e^{-\Lambda t}\end{aligned}$$

suggesting a mixing time of order x starting from x. This is ok for small x. Fixing $x_0 \in \mathbb{Z}_+$ to be specified later on, it leads us to consider

$$S^x := \inf\{t \geqslant 0 : X_t^x \leqslant x_0\}$$

Lemma 9

There exist (random) constants $x_0 \in \mathbb{Z}_+$ and $a > 0$ such that

$$\forall\, x \in \mathbb{Z}_+, \qquad \mathbb{E}[S^x] \leqslant 2\log_a^*(x)$$

Proof of Theorem 1

Let be given $t \geqslant 0$ and $A \subset \mathbb{Z}_+$. Conditioning by the events before S^x, say $\mathcal{F}_{S^x}$, we get for any $x \in \mathbb{Z}_+$,

$$
\begin{aligned}
|\mathcal{L}(X_t^x)[A] - \pi(A)| &= |\mathbb{P}[X_t^x \in A] - \pi(A)| \\
&\leqslant |\mathbb{P}[X_t^x \in A|\mathcal{F}_{S^x}, S^x \leqslant t/2] - \pi(A)| \\
&\quad + \mathbb{P}[S^x > t/2]
\end{aligned}
$$

On one hand, using the strong Markov property at S^x and the bound deduced from the spectral gap, and on the other hand resorting to Lemma 9, we get

$$
|\mathcal{L}(X_t^x)[A] - \pi(A)| \leqslant \frac{t}{2}(x_0+1)^K 2^{x_0/2} e^{-\Lambda t/2} + \frac{4}{t}\log_a^*(x)
$$

and the r.h.s. can be made as small as we desire, and uniformly in A, by the choice of $t = b(1+\log_a^*(x))$ and b large enough. ∎

Proof of Lemma 9 (1)

It is sufficient to work with the imbedded Markov chain, abusively denoted the same. Recalling Lemma 4, define

$$\forall\ n \in \mathbb{Z}_+, \qquad Z_n^x \quad := \quad \log_a^*(X_n^x)$$

with $a := 2^{\frac{1}{2K}}$.
Lemma 9 will be proven if we can find x_0 (independent from x) such that

$$\left(Z_{n\wedge S^x}^x + \frac{n \wedge S^x}{2} \right)_n$$

is a supermartingale. Indeed, letting n go to infinity in

$$\mathbb{E}\left[Z_{n\wedge S^x}^x + \frac{n \wedge S^x}{2} \right] \quad \leqslant \quad Z_0^x \quad = \quad \log_a^*(x)$$

we get the desired bound.

The previous supermartingale property amounts to see that for any $x \geqslant x_0$,

$$\mathbb{E}[Z_1^x - Z_0^x] \leqslant \frac{1}{2} \tag{4}$$

Indeed, consider $y := 2K \log(x) = \log_a(x)$. For $z \leqslant y$, we have

$$\log_a^*(z) \leqslant \log_a^*(y) = \log_a^*(\log_a(x)) = \log_a^*(x) - 1$$

thus

$$\mathbb{E}[(Z_1^x - Z_0^x)\mathbb{1}_{X_1^x \leqslant y}] \leqslant (-1)\mathbb{P}[X_1^x \leqslant y] = -1 + \mathbb{P}[X_1^x > y]$$

Proof of Lemma 9 (3)

We deduce, via Lemma 4,

$$\begin{aligned}
\mathbb{E}[Z_1^x - Z_0^x] &\leqslant -1 + \sum_{z>y} (1 + \log_a^*(z)) \frac{Q(z)}{Q(N(x))} \\
&\leqslant -1 + \frac{1}{Q(p(x))} \sum_{z>y} (1 + \log_a^*(z)) Q(z) \\
&\leqslant -1 + \frac{1}{Q(\lfloor K \log(x) \rfloor)} \sum_{z>y} (1 + \log_a^*(z)) Q(z) \\
&= -1 + \frac{1}{Q(\lfloor y/2 \rfloor)} \sum_{z>y} (1 + \log_a^*(z)) Q(z)
\end{aligned}$$

Due to the exponential feature of Q, the last term of the r.h.s. is as small we desire for y large enough. We can thus find $x_0 \in \mathbb{Z}_+$, such that (4) holds for any $x \geqslant x_0$.

$\blacksquare$

Plan

A family of subtrees

Since the above approach is based on the analysis of the smallest Dirichlet eigenvalue on trees whose root is absorbing, we can also resort to Hardy's inequalities.
Let us recall the principle. See $\mathcal{T}$ as a tree rooted in 0 and for any $x \in \mathbb{Z}_+$, denote by $h(x)$ the height of x in $\mathcal{T}$. Consider $\mathbb{T}$ the set of all subtrees T of $\mathcal{T}$ satisfying the conditions

- T does not contain 0,
- there exists $M \geqslant 1$ such that $h(x) \leqslant M$ for all $x \in T$,
- if $x \in T$ has a child in T, then all children of x belong to T.

Each $T \in \mathbb{T}$ has a root that is denoted $r(T)$. When T is not reduced to a singleton, the set of sons of $r(T)$ will be denoted $\mathcal{S}(T)$ (it is the same in $\mathcal{T}$ and in T). For $y \in \mathcal{S}(T)$, write T_y for the set of all offsprings of y in T, so we have the decomposition

$$T \quad = \quad \{r(T)\} \sqcup \bigsqcup_{y \in \mathcal{S}(T)} T_y$$

Recall we have defined a functional on the edges:

$$\nu(\{x,y\}) \quad := \quad Z\pi(x)K_{\mathcal{T}}(x,y)$$

(where Z is the normalization in π). We now extend it on $\mathbb{T}$ (no longer as a measure) via the iteration

- when T is the singleton $\{r(T)\}$, we take

$$\nu(T) := \nu(\{r(T), p(r(T))\})$$

- when T is not a singleton, then ν satisfies

$$\frac{1}{\nu(T)} \quad = \quad \frac{1}{\nu(\{r(T)\})} + \frac{1}{\sum_{y\in\mathcal{S}(T)} \nu(T_y)}$$

Also for $T \in \mathbb{T}$, let T^* be the set of all offsprings in $\mathcal{T}$ of the leaves of T (themselves included).

Consider $\mathbb{S} \subset \mathbb{T}$ the set of $T \in \mathbb{T}$ which are such that $r(T)$ is a son of 0.
Finally define

$$A \quad := \quad \sup_{T \in \mathbb{S}} \frac{\mu(T^*)}{\nu(T)}$$

The interest of this quantity is the Hardy inequality:

$$A \; \leqslant \; \frac{1}{\lambda} \; \leqslant \; 16A \tag{5}$$

where λ is the best constant in Proposition 7, namely the smallest Dirichlet eigenvalue for the Markov process associated to $K_{\mathcal{T}}$ and absorbed at 0.

An alternative proof of the spectral gap

It is sufficient to show that $A < +\infty$ (a.s.).
It is possible to do so, based on Lemmas 2, 3 and 4. This proof is more involved than the use of the Cheeger's inequality. But a priori it provides a better estimate, since (5) gives almost matching upper and lower bounds.

To illustrate the improvement, let us turn to a deterministic situation.

Plan

On the set of vertices $\mathbb{Z}_+$, an edge is put between $x < y$ when in the dyadic decomposition of y, the coefficient in front of 2^x is 1. Exemple: $9 = 2^0 + 2^3$, so the edges $\{0, 9\}$ and $\{3, 9\}$ belong to the Rado graph, called G.
Thus for any given $x \in \mathbb{Z}_+$, the neighbors $y > x$ are exactly the $n2^x$, where n is an odd number.

The graph G is isomorphic to $\mathcal{G}$, a.s.

All the previous notions, such as the set of neighbors $N(x)$ or the parent $p(x)$ of a vertex $x \in \mathbb{Z}_+$, are defined as before.

For any $\delta \in (0,1)$, let Q be the probability given by

$$\forall\ x \in \mathbb{Z}_+, \qquad Q(x) \quad := \quad (1-\delta)\delta^x$$

and consider the associated sphere random walk.
The previous arguments can be extended (simplified in fact) to get

Theorem 10

There exists a constant $b > 0$ depending on δ such that for any $x \in \mathbb{Z}_+$, the mixing time τ^x satisfies

$$\tau^x \quad \leqslant \quad b(1 + \log_2^*(x))$$

Cheeger's vs Hardy's bounds

Concerning the proof of Proposition 7, giving an estimate of a smallest Dirichlet eigenvalue on a tree rooted in 0, both Cheeger or Hardy methods are available. Cheeger's inequality leads to the bound

$$\lambda \geqslant \frac{(1-\delta)^2}{2} \tag{6}$$

while Hardy's inequality implies

$$\lambda \geqslant \frac{1-\delta}{16(2 \vee \lceil \log\log(2/\log(1/\delta))\rceil)}$$

which is better than (6) as δ goes to $1-$ ($\delta = 1$ would correspond to the problematic case "pick a neighbor at random").

Questions

The previous models lead to a lot of unanswered questions. Here are two examples for the determinist Rado graph.

- It can be shown the smallest Dirichlet eigenvalue of Proposition 7 is positive if and only if $Q(x)$ is of the same order as $Q([\![x,\infty[\![)$. Is this assertion also true for the spectral gap of the corresponding sphere random walk? More generally and in the spirit of [Benjamini and Schramm, Every graph with a positive Cheeger constant contains a tree with a positive Cheeger constant, 1997], does a Markov process with a positive Cheeger constant "contain" a Markov process induced on a tree with a positive Cheeger constant?
- For the sphere Markov process on the Rado graph, do the pruning procedure and the Hardy's estimate give the right order of the spectral gap?