

## A product chain without cutoff

Hubert Lacoin\*

### Abstract

In this note, we construct an example of a sequence of  $n$ -fold product chains which does display cutoff neither for the total-variation distance nor for the separation distance. In addition we show that this type of product chains necessarily displays pre-cutoff.

**Keywords:** Markov chain; cutoff; counter-example.

**AMS MSC 2010:** 60J10.

Submitted to ECP on September 1, 2014, final version accepted on January 12, 2015.

Supersedes arXiv:1407.1754v2.

## 1 Introduction

Consider a sequence of reversible irreducible continuous Markov chains  $X^n = (X^n(t))_{t \geq 0}$ , each being defined on a finite state spaces  $(\Omega_n)_{n \geq 0}$ . Let  $\pi_n$  denote the unique reversible probability measure associated to  $X^n$ . It is a classic result of Markov chain theory that for any initial condition the distribution of  $X^n(t)$  converges to  $\pi_n$  when  $t$  goes to infinity. We let  $P_t^n$  denote the Markov semigroup associated to  $X^n$  and  $d_n(t)$  resp.  $d_n^s(t)$  denote the distance to equilibrium for the total variation distance and separation distance (they are defined by taking the maximal distance over all initial condition)

$$\begin{aligned} d_n(t) &:= \max_{x \in \Omega_n} \|P_t^n(x, \cdot) - \pi_n\|_{TV}, \\ d_n^s(t) &:= 1 - \min_{x, y \in \Omega_n} \frac{P_t^n(x, y)}{\pi_n(y)}. \end{aligned} \tag{1.1}$$

When we have to consider only one Markov chain  $X$  in Section 2.2, we will use the same notation without  $n$ .

The sequence  $X^n$  is said to display *cutoff* if  $d_n(t)$  drops abruptly from 1 to 0 on the appropriate time scale. More precisely, if one defines the mixing time corresponding to the distance  $a \in (0, 1)$  to be

$$t_{\text{mix}}^n(a) := \inf\{t \mid d_n(t) < a\}. \tag{1.2}$$

the chain is said to display *cutoff* if for any  $\varepsilon \in (0, 1/2]$

$$\lim_{n \rightarrow \infty} t_{\text{mix}}^n(\varepsilon)/t_{\text{mix}}^n(1 - \varepsilon) = 1. \tag{1.3}$$

---

\*IMPA, Rio de Janeiro, Brazil. E-mail: lacoin@impa.br

We follow the definition given in [5, pp .248] and say displays pre-cutoff if

$$\limsup_{\varepsilon \rightarrow 0+} \limsup_{n \rightarrow \infty} t_{\text{mix}}^n(\varepsilon)/t_{\text{mix}}^n(1 - \varepsilon) < \infty. \quad (1.4)$$

Note that one can replace  $t_{\text{mix}}^n$  by  $t_s^n$  the mixing time for the separation distance.

The term cutoff was coined by Aldous and Diaconis [1] and its occurrence for the transposition shuffle was proved by Diaconis and Shahshahani [4]. It is thought to hold for many natural sequences of Markov chain as soon as

$$t_{\text{mix}}^n(1/4) \times \text{gap}_n = \infty \quad (H)$$

where  $\text{gap}_n$  corresponds to the spectral gap of the chain  $X^n$  (see e.g. [5, Chapter 12 and Chapter 18] for the definition of the spectral gap and an account on the cutoff phenomenon). More precisely the condition is necessary and it was proposed by Peres as a natural sufficient condition provided the chain is “nice enough”. As (H) is in fact known to be a necessary condition for pre-cutoff, this would imply in particular that pre-cutoff implies cutoff for “nice chains”.

Shortly after (H) was proposed as a sufficient condition for cutoff, Aldous constructed a chain that satisfies (H) and displays pre-cutoff, but for which cutoff does not hold. Pak also constructed a counter-example (with no pre-cutoff) which is a random walk on a Cayley graph (see [5, pp 253–256]). Since then it has been a challenge to find a large class of Markov chain for which the (H) condition is a sufficient one. Note that Chen and Saloff-Coste have shown that (H) is a sufficient condition in full generality when distance to equilibrium is measured by the  $L^p$  norm [3]. Let us note also that [6, Proposition 7] establishes that cutoff holds for large product chains provided one has a good-control on the supremum norm of the relative density of the marginals.

We define  $Y^n$  the chain corresponding to  $n$  independent copies of  $X^n$  (its  $n$ -th power)

$$Y^n(t) := (X_1^n(t), \dots, X_n^n(t)). \quad (1.5)$$

In this note we show that the sequence  $Y^n$  always displays pre-cutoff, and we construct a sequence of chain  $X$  which is such that  $Y$  displays no cutoff (whereas  $X$  does), showing that condition (H) is not a sufficient condition for cutoff for chains that are large powers of a simpler one.

## 2 Pre-cutoff for product chains

We let  $D_n$ ,  $D_n^s$ ,  $Q_t^n$ ,  $T_{\text{mix}}^n$  and  $T_s^n$ , and  $\mu_n := \pi_n^{\otimes n}$  denote the distances to equilibrium, semigroup, and mixing time and equilibrium measure for the chain  $Y^n$ . We have

**Proposition 2.1.** *For any sequence of non-trivial Markov chain  $X^n$  one has*

$$\limsup_{n \rightarrow \infty} \frac{T_{\text{mix}}^n(1 - \varepsilon)}{T_{\text{mix}}^n(\varepsilon)} \leq 2. \quad (2.1)$$

*The result also holds when the total-variation distance is replaced by the separation distance.*

**Remark 2.2.** In the first draft of this paper, the optimal bound of 2 for the mixing time ratio was proved to hold only for the separation distance. The idea of using the Hellinger distance to obtain an optimal bound also for the total-variation distance (developped in Section 2.2) is due to Yuval Peres.

**2.1 Proof of Proposition 2.1 for the separation distance**

The separation distance to equilibrium for  $Y^n$  is given by

$$D_n^s(t) := 1 - \min_{\mathbf{x}, \mathbf{y} \in \Omega_n} \frac{Q_t^n(\mathbf{x}, \mathbf{y})}{\mu_n(\mathbf{y})} = 1 - (1 - d_n^s(t))^n. \tag{2.2}$$

Hence for  $\varepsilon \in (0, 1/2)$  fixed and  $n$  sufficiently large we have

$$t_s^n(n^{-2/3}) \leq T_s^n(1 - \varepsilon) \leq T_s^n(\varepsilon) \leq t_s^n(n^{-4/3}) \leq 2t_s^n(n^{-2/3}), \tag{2.3}$$

where the last inequality is due to the submultiplicativity property for the separation distance

$$d_n^s(a + b) \leq d_n^s(a)d_n^s(b). \tag{2.4}$$

Hence the result. □

For the total-variation distance, can obtain (2.1) with 4 instead of 2 on the r.h.s. simply by using the following comparison between the total variation distance and separation distance for reversible Markov chains initially proved in [2] (see also [5, Lemma 6.13 and Lemma 19.3])

$$d_n(t) \leq d_n^s(t) \leq 4d_n(t/2).$$

**2.2 Proof of Proposition 2.1 for the total-variation distance**

For an optimal result, we need to use the *Hellinger distance* which has the property of behaving nicely for product. This section starts with the introduction of notation and recalling some classical inequalities.

Given  $\mu$  and  $\nu$  two probability measures on a common finite state space  $\Omega$ , the Hellinger distance between  $\mu$  and  $\nu$  is defined by

$$d^H(\mu, \nu) := \sqrt{\sum_{x \in \Omega} (\sqrt{\nu(x)} - \sqrt{\mu(x)})^2}. \tag{2.5}$$

We have the following comparisons with the total-variation distance (see for instance [5, (20.22) and (20.29)])

$$\|\mu - \nu\|_{TV} \leq d^H(\mu, \nu) \leq \sqrt{2\|\mu - \nu\|_{TV}} \tag{2.6}$$

We set

$$d_n^H(t) := \sup_{x \in \Omega} d^H(P_t^n(x, \cdot), \pi_n), \tag{2.7}$$

and let  $D_n^H(t)$  denote the counterpart of  $d_n^H$  for the chain  $Y_n$ . Similarly to (2.2), it is easy to remark (see also [5, Exercice 20.5]) that

$$1 - \frac{1}{2} (D_n^H(t))^2 = \left(1 - \frac{1}{2} (d_n^H(t))^2\right)^n. \tag{2.8}$$

Hence we know that  $D_n^H(t)$  is close to  $\sqrt{2}$  resp. 0 (and hence by (2.6) that  $D_n(t)$  is close to 1 resp. 0) if and only if  $\sqrt{n}d_n^H(t)$  is close to infinity resp. 0. What we need to conclude is that there is a time window  $[t_n, 2t_n]$  for which the Hellinger distance drops from  $n^{-1/2+\delta}$  to  $n^{-1/2-\delta}$ . We achieve this by proving the following property of the Hellinger distance for reversible Markov chains

**Lemma 2.3.** *For any reversible irreducible Markov chain and any  $t \geq 0$ ,*

$$d^H(2t) \leq 8(d^H(t))^{5/4}. \tag{2.9}$$

With this results at hand, it is easy to prove, that similarly to (2.3), for

$$t_n := \inf\{t \mid d_n^H(t) \leq n^{-3/7}\}.$$

one has for any  $\varepsilon \in (0, 1/2)$ , for all  $n$  sufficiently large

$$t_n \leq T_{\text{mix}}^n(1 - \varepsilon) \leq T_{\text{mix}}^n(\varepsilon) \leq 2t_n. \tag{2.10}$$

*Proof of Lemma 2.3.* We introduce now  $\bar{d}(t)$  defined as

$$\bar{d}(t) := \max_{x,y \in \Omega^2} \|P_t(x, \cdot) - P_t(y, \cdot)\|_{TV}. \tag{2.11}$$

Note that, as the chain is assumed to be reversible  $\bar{d}(t)$  also correspond to the operator norm for  $P_t$  acting on integrable functions with mean 0, or more precisely

$$\bar{d}(t) = \max_{\{f \in L_1(\pi) \mid \pi(f)=0\}} \frac{\|P_t f\|_{L_1(\pi)}}{\|f\|_{L_1(\pi)}}, \tag{2.12}$$

where

$$P_t f(x) := \sum_{y \in \Omega} P_t(x, y) f(y).$$

The function  $\bar{d}(t)$  compares well with  $d(t)$  and is submultiplicative (see for instance [5, Chapter 4])

$$\begin{aligned} d(t) &\leq \bar{d}(t) \leq 2d(t) \\ \bar{d}(t+s) &\leq \bar{d}(t)\bar{d}(s). \end{aligned} \tag{2.13}$$

Combining (2.6) and (2.13), we have for every  $t$

$$\bar{d}(t)/2 \leq d^H(t) \leq \sqrt{2\bar{d}(t)}. \tag{2.14}$$

Let us try now to prove the result from (2.14) (inequality on the left) and (2.13) in a naive way. We have

$$\bar{d}(2t) \leq (\bar{d}(t))^2 \leq 4(d^H(t))^2, \tag{2.15}$$

and hence using (2.14) again (inequality on the right) we obtain

$$d^H(2t) \leq \sqrt{8}d^H(t), \tag{2.16}$$

which is not satisfying.

To find a way out, we have to prove that if the inequality on the left in (2.14) is sharp for  $t$ , the inequality on the right cannot be sharp for  $2t$ .

We set  $u := d^H(t)$  (note that we can assume  $u \leq 1$  as the result is trivial for  $u \geq 1$ ) Let  $x$  an element of  $\Omega$  for which  $d^H(2t) = d^H(P_t(x, \cdot), \pi)$ . Let  $g$  denote the density of  $P_t(x, \cdot)$  with respect to  $\pi$  and  $g'$  denote the density of  $P_{2t}(x, \cdot)$  with respect to  $\pi$ .

We have from our definitions

$$\begin{aligned} \int \left(\sqrt{g'(y)} - 1\right)^2 \pi(\mathrm{d}y) &= (d_H(2t))^2, \\ \int \left(\sqrt{g(y)} - 1\right)^2 \pi(\mathrm{d}y) &\leq u^2. \end{aligned} \tag{2.17}$$

Our first step is the contribution to the total variation distance  $\|P_t(x, \cdot) - \pi\|$  of the set  $\{y \mid |g(y) - 1| \geq u^{1/2}\}$  is much smaller than  $u$ .

**Lemma 2.4.** We have for all  $u \leq 1$

$$\int |g(y) - 1| \mathbf{1}_{\{|g(y)-1| \geq u^{1/2}\}} \, d\pi(\mathbf{d}y) \leq 10u^{3/2}. \quad (2.18)$$

*Proof.* We have to show that

$$\begin{aligned} \int |g - 1| \mathbf{1}_{\{|g(y)-1| \geq u^{1/2}\}} \, d\pi(\mathbf{d}y) \\ \leq 10 \int \left( \sqrt{g(y)} - 1 \right)^2 u^{-1/2} \mathbf{1}_{\{|g(y)-1| \geq u^{1/2}\}} \, d\pi(\mathbf{d}y), \end{aligned} \quad (2.19)$$

and we conclude by using (2.17). The inequality (2.19) is obtained by noticing that when  $g \geq 2$  we have

$$|g - 1| \leq (3 - 2\sqrt{2})|\sqrt{g} - 1|^2, \quad (2.20)$$

while when  $g \in (0, 2)$ ,  $|g - 1| \geq u^{1/2}$ , we have

$$|g - 1| \leq u^{-1/2}|g - 1|^2 \leq \frac{u^{-1/2}}{(\sqrt{2} - 1)^2} |\sqrt{g} - 1|^2. \quad (2.21)$$

□

Now we can decompose  $g - 1$  into a sum of two function  $h_1$  and  $h_2$ : one which has a small  $l_\infty$  norm, and one which has a small  $l_1$  norm.

$$\begin{aligned} h_1(y) &:= (g - 1)(y) \mathbf{1}_{\{|g(y)-1| < u^{1/2}\}}, \\ h_2(y) &:= (g - 1)(y) \mathbf{1}_{\{|g(y)-1| \geq u^{1/2}\}}. \end{aligned} \quad (2.22)$$

We have

$$\begin{aligned} \|h_1\|_{l_\infty} &\leq u^{1/2}, \\ \|h_2\|_{l_1(\pi)} &\leq 10u^{3/2}. \end{aligned} \quad (2.23)$$

Setting  $h'_i := P_t h_i$  one has

$$g' - 1 = h'_1 + h'_2. \quad (2.24)$$

From (2.12) one has (using (2.14) to bound  $\bar{d}(t)$ )

$$\begin{aligned} \|h'_2\|_{l_1(\pi)} &\leq \bar{d}(t) \|h_2\|_{l_1(\pi)} \leq 20u^{5/2}, \\ \|h'_1\|_{l_\infty} &\leq \|h_1\|_{l_\infty} \leq u^{1/2}. \end{aligned} \quad (2.25)$$

Moreover

$$\|g' - 1\|_{l_1(\pi)} \leq 2d(2t) \leq 2\bar{d}(t)^2 \leq 8u^2. \quad (2.26)$$

We are now ready to bound  $(d^H(2t))^2$ . We split it into two parts. The first one is bounded thanks to (2.26)

$$\int \left( \sqrt{g'(y)} - 1 \right)^2 \mathbf{1}_{\{|g'(y)-1| \leq 2u^{1/2}\}} \pi(\mathbf{d}y) \leq 2u^{1/2} \int |g'(y) - 1| \pi(\mathbf{d}y) \leq 16u^{5/2}. \quad (2.27)$$

For the second part, note that as  $(\sqrt{g'} - 1)^2 \leq |g' - 1|$  we have

$$\begin{aligned} \int \left( \sqrt{g'(y)} - 1 \right)^2 \mathbf{1}_{\{|g'(y)-1| \geq 2u^{1/2}\}} \pi(\mathbf{d}y) \\ \leq \int |g'(y) - 1| \mathbf{1}_{\{|h'_2(y)| \geq u^{1/2}\}} \pi(\mathbf{d}y) \\ \leq \int 2|h'_2(y)| \mathbf{1}_{\{|h'_2(y)| \geq u^{1/2}\}} \pi(\mathbf{d}y) \leq 40u^{5/2}. \end{aligned} \quad (2.28)$$

where the last inequality comes from (2.25), and the one before from the fact that

$$|g' - 1| \leq |h'_1 + h'_2| \leq |h'_2| + u^{1/2}.$$

This allows us to conclude. □

### 3 An example without cutoff

#### 3.1 Construction

Let us now define a sequence  $X^n$  such that  $Y^n$  displays no cutoff. The idea build on the counter example of Aldous displayed on [5, pp 256]. The state-space of  $X^n$  is the vertex set  $V_n$  of a graph  $G_n$  with  $2n + 1$  edges and  $2n + 1$  vertices defined as follows:

- There is a segment of  $2n$  edges linking  $2n + 1$  vertices. We call  $A$  and  $C$  its ends.
- There is an extra edge linking the middle point of the segment (which we call  $B$ ) to  $C$ .

The transition rates are positive on the edges of  $G_n$  and are specified in the caption of Figure 1 .

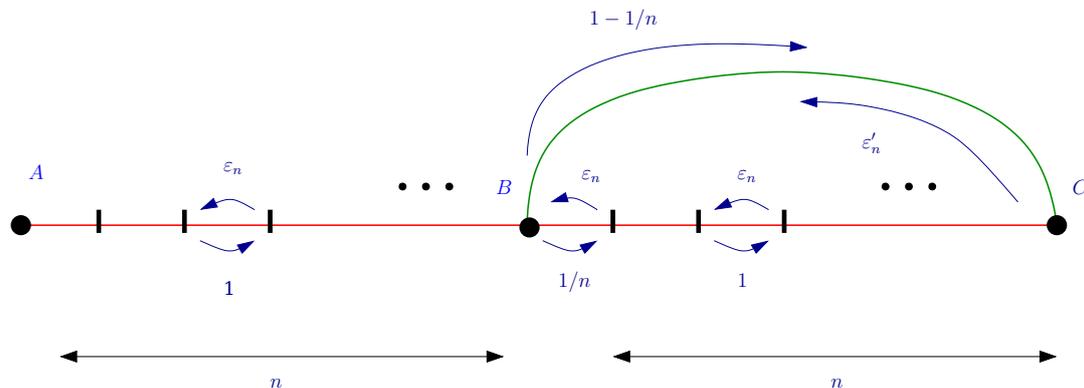


Figure 1: The graph  $G_n$  together with the transition rates of  $X^n$ : the two segments that are represented in red are of length  $n$ . The jump rate are represented in blue above the arrows. In the direction from  $A$  to  $C$  the jump rate is always one except at point  $B$  where the jump rate to  $C$  (along the green edge) is equal to  $1 - 1/n$  while the probability to jump towards  $C$  on the red path is  $1/n$ . The jumps in the direction of  $A$  along red edges are equal to  $\varepsilon_n := 2^{-n^2}$ . The jump rate from  $C$  to  $B$  along the green edge is equal to  $\varepsilon'_n := (n - 1)2^{-n^3}$ .

With this definition it is not difficult to check that  $X^n$  is a reversible Markov chain. We have chosen  $\varepsilon_n$  to be exponentially small but the result we are going to present would remain valid for  $\varepsilon_n = 1/2$  for all  $n$  (or any other value smaller than 1). Note that the value of  $\varepsilon'_n$  is determined by that of  $\varepsilon_n$  in order to have reversibility.

**Proposition 3.1.** *The construction above satisfies the following property*

- (i) *The sequence  $X^n$  displays cutoff around time  $n$ , both in separation and total-variation distance.*
- (ii) *The sequence does not  $Y^n$  display cutoff as*

$$T_{\text{mix}}^n(a) = \begin{cases} 2n(1 + o(1)) & \text{for } a \in (0, 1 - e^{-1}), \\ n(1 + o(1)) & \text{for } a \in (1 - e^{-1}, 1). \end{cases} \quad (3.1)$$

(the notation means that for a fixed  $a \neq (1 - e^{-1})$ ,  $T_{\text{mix}}^n(a)/n$  converges either to 1 or 2.) The same holds for the separation distance.

**Remark 3.2.** The above Proposition shows that the inequality (2.1) concerning the ratio of the mixing time is optimal.

The main idea of the proof is that the total variation distance can be expressed in terms of the distribution of the time  $\tau$  or  $\mathcal{T}$  needed to reach  $A$  (for  $X^n$ ) or  $\mathbf{C} := (C, C, \dots, C)$  (for  $Y^n$ ) starting from  $A$ . In particular, there is cutoff if and only if this time is concentrated around its mean. For  $X^n$  we show that  $\tau$  concentrates around  $n$ , whereas for  $Y^n$ ,  $\mathcal{T}$  will be about  $2n$  if at least one of the coordinates  $X_1^n$  decides to use the red path between  $B$  and  $C$  (which happens with a non-vanishing probability).

### 3.2 Proof of Proposition 3.1

The equilibrium measure  $\pi_n$  gives a weight  $1 - O(2^{-n^2})$  to the vertex  $C$ , and hence the equilibrium measure  $\mu_n$  of  $Y^n$ , gives weight  $1 - O(n2^{-n^2})$  to  $\mathbf{C} := (C, C, \dots, C)$ . Because of this remark we have

$$d_n(t) = 1 - \min_{x \in V_n} P_t(x, C) + o(1) \quad \text{and} \quad D_n(t) = 1 - \min_{\mathbf{x} \in V_n^n} P_t(\mathbf{x}, \mathbf{C}) + o(1). \quad (3.2)$$

For  $x \in V_n$  or  $\mathbf{x} \in V_n^n$ , let  $\mathbb{P}^{n,x}$  resp.  $\mathbb{Q}^{n,\mathbf{x}}$  be the law of  $X^n(t)$  starting from  $x$  resp. the law of  $Y^n(t)$  and let  $\tau$ , resp.  $\mathcal{T}$  be the first hitting time of  $C$  resp.  $\mathbf{C}$ .

**Lemma 3.3.** We have

$$\begin{aligned} d_n(t) &= \mathbb{P}^{n,A}(\tau > t) + o(1), \\ D_n(t) &= \mathbb{Q}^{n,\mathbf{A}}(\mathcal{T} > t) + o(1), \end{aligned} \quad (3.3)$$

meaning that

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{t \geq 0} |d_n(t) - \mathbb{P}^{n,A}(\tau > t)| &= 0, \\ \lim_{n \rightarrow \infty} \sup_{t \geq 0} |D_n(t) - \mathbb{Q}^{n,\mathbf{A}}(\mathcal{T} > t)| &= 0. \end{aligned} \quad (3.4)$$

*Proof.* We provide the proof for  $d_n(t)$  as the other is identical. First let us prove the result for  $t < 3n$ , and we will check later that for  $t > 3n$  both  $d_n(t)$  and  $\mathbb{P}^n(\tau > t)$  are  $o(1)$ . Now the probability that a jump in the direction  $A$  (a backtrack) occurs before time  $3n$  is exponentially small in  $n$  and thus from (3.2) we have

$$P_t(x, C) = \mathbb{P}^{n,x}(\tau \leq t) + o(1) \quad (3.5)$$

Hence from (3.2), it is sufficient to check that the minimum of  $\mathbb{P}^{n,x}(\tau \leq t)$  is reached for  $A$  (up to some  $o(1)$  correction).

From an obvious coupling, we see that  $A$  is the point of the segment  $AB$  which makes  $\tau$  the largest. It remains to check that starting from one of the  $n - 1$  inside points the red segment  $B$  and  $C$  cannot make  $\tau$  larger: by conditioning to the event that  $X^n$  does not backtrack before  $t$  (which is an event of almost full probability) we see that  $\tau$  starting from  $A$  is bounded from below by a sum of  $n + 1$  IID standard exponentials whereas in the  $BC$  branches it is bounded from above by the sum of  $n$  IID standard exponentials.

Finally, for  $t = 3n$ , as conditioned on no backtrack,  $\tau$  starting from  $A$  is a bounded from above by a sum of  $2n$  IID standard exponentials, both  $\mathbb{P}^n(\tau > 3n)$  and  $d_n(t)$  are  $o(1)$  (and the fact both functions are decreasing allows to conclude for larger values of  $t$ ).  $\square$

From Lemma (3.3) one has

$$D_n(t) = 1 - [\mathbb{P}^{n,A}(\tau \leq t)]^n + o(1). \tag{3.6}$$

Hence  $D_n(t)$  is in a neighborhood of 1 resp. 0 if and only if  $n\mathbb{P}^{n,A}(\tau > t)$  is in a neighborhood of infinity resp. 0.

Concerning  $X^n$ , one can remark that conditioning to the event that  $X^n$  does not backtrack and uses a short branch to reach  $C$ ,  $\tau$  is a sum of  $3n + 1$  IID standard exponentials. Hence as the event to which we are conditioning has a probability tending to one, we have

$$\lim_{n \rightarrow \infty} d_n(ns) = \begin{cases} 1 & \text{if } s < 1, \\ 0 & \text{if } s > 1. \end{cases} \tag{3.7}$$

and  $X^n$  exhibits cutoff. However, the slow branch plays a crucial role for the product chain as the probability to hit  $C$  from the longer branch asymptotically behaves like  $n^{-1}$ . As a consequence we have

**Lemma 3.4.**

$$\lim_{n \rightarrow \infty} n\mathbb{P}^{n,A}(\tau > ns) = \begin{cases} \infty & \text{if } s < 1, \\ 1 & \text{if } s \in (1, 2), \\ 0 & \text{if } s > 2. \end{cases} \tag{3.8}$$

*Proof.* Under  $\mathbb{P}^{n,A}$  the probability that  $X^n$  backtrack before time  $3n$  is exponentially small in  $n$  and thus can be neglected. Conditioned on no backtracking, the probability to use the red segment  $BC$  is equal to  $n^{-1}$ . Now conditioned on using the red segment,  $\tau$  is a sum of  $2n$  IID standard exponentials whereas conditioned on using the green edge  $\tau$  is a sum of  $n + 1$  IID standard exponentials. Hence the result.  $\square$

This implies

$$\lim_{n \rightarrow \infty} D_n(ns) = \begin{cases} 1 & \text{if } s < 1, \\ 1 - e^{-1} & \text{if } s \in (1, 2), \\ 0 & \text{if } s > 2. \end{cases} \tag{3.9}$$

and hence  $Y^n$  exhibits no cutoff for total variation distance.

Now let us show that cutoff also holds for the separation distance. This amounts essentially to prove the following

**Lemma 3.5.** *For all  $n$  sufficiently large, for any  $x, y \in V_n \setminus \{C\}$ , for all  $n/2 \leq t \leq 3n$  one has*

$$P_t^n(x, y) \geq \pi_n(y) \tag{3.10}$$

*Proof.* From reversibility

$$\frac{P_t^n(x, y)}{\pi_n(y)} = \frac{P_t^n(y, x)}{\pi_n(x)} \tag{3.11}$$

so that one can without loss of generality consider that  $x$  is the point closer to  $A$  on the red segment. Let  $d$  be the number of red edges between  $x$  and  $y$ . Then  $P_t^n(x, y)$  is bounded from below by the probability of the event: in the time interval  $[0, t]$  the walk  $X^n$  (starting from  $x$ ) makes exactly  $d$  jumps following the red path from  $x$  to  $y$ .

As the jump rate for  $X$  is always of order 1 (except at point  $C$ ), the probability of making exactly  $d$  jumps in the time interval  $[0, t]$  is larger than  $e^{-C_1 n}$  for some constant

$C_1$ . The probability of not following the red path conditioning to the number of jump is at least  $1/2n$  (a backtrack is exponentially unlikely, and if the path goes through  $B$  the chance of choosing the right direction there is equivalent to  $n^{-1}$ ). Hence there exists a constant  $C_2$  such that when  $n$  is sufficiently large

$$\forall t \in (n/2, 3n), \quad \forall x, y \in V_n \setminus \{C\}, P_t^n(x, y) \geq e^{-C_2 n}. \quad (3.12)$$

As  $\pi_n(y) \leq 2^{-n^2}$  for all  $y \neq C$ , this is sufficient to conclude.  $\square$

From the previous Lemma (and the definition (1.1) and reversibility), one has for all  $t \in (n/2, 3n)$

$$d_n^s(t) := 1 - \min_{x \in \Omega_n} \frac{P_t^n(x, C)}{\pi_n(C)}, \quad (3.13)$$

which according to (3.2) shows that the difference between total-variation and separation distance for this chain is negligible.

## References

- [1] D. Aldous and P. Diaconis. *Shuffling Cards and Stopping Times* The American Mathematical Monthly, **93** (1986) 333-348. MR-0841111
- [2] D. Aldous and P. Diaconis *Strong uniform times and finite random walks* Adv. in Appl. Math. **8** (1987) 69-97. MR-0876954
- [3] G. Chen and L. Saloff-Costes. *The Cutoff Phenomenon for Ergodic Markov Processes*, Elec. J. Probab. **13** (2008) 26-76. MR-2375599
- [4] P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions* Z. Wahrsch. Verw. Gebiete **57** (1981) 159-179. MR-0626813
- [5] D. Levin, Y. Peres, E. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, Providence, RI, (2009). MR-2466937
- [6] E. Lubetzky and A. Sly *Cutoff for General Spin Systems with Arbitrary Boundary Conditions* Comm. Pure and Appl. Math. **67** (2014) 982-1027. MR-3193965

**Acknowledgments.** The author is grateful to Perla Sousi and Yuval Peres to have make him now about the question of cutoff for product chains and for enlightening discussions. In particular the author wishes to thank Yuval Peres for suggesting a proof of (2.1) for the total-variation distance.

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

### Economical model of EJP-ECP

- Low cost, based on free software (OJS<sup>1</sup>)
- Non profit, sponsored by IMS<sup>2</sup>, BS<sup>3</sup>, PKP<sup>4</sup>
- Purely electronic and secure (LOCKSS<sup>5</sup>)

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

<sup>2</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>3</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>4</sup>PK: Public Knowledge Project <http://pkp.sfu.ca/>

<sup>5</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>