



HAL
open science

Positive association of the oriented percolation cluster in randomly oriented graphs

François Bienvenu

► **To cite this version:**

François Bienvenu. Positive association of the oriented percolation cluster in randomly oriented graphs. 2017. hal-01649325v1

HAL Id: hal-01649325

<https://hal.science/hal-01649325v1>

Preprint submitted on 27 Nov 2017 (v1), last revised 16 Feb 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Positive association of the oriented percolation cluster in randomly oriented graphs

François Bienvenu ^{*†}

November 23, 2017

Abstract

Consider any fixed graph whose edges have been randomly and independently oriented, and write $\{S \rightsquigarrow i\}$ to indicate that there is an oriented path going from a vertex $s \in S$ to vertex i . Narayanan (2016) proved that for any set S and any two vertices i and j , $\{S \rightsquigarrow i\}$ and $\{S \rightsquigarrow j\}$ are positively correlated. His proof relies on the Ahlswede-Daykin inequality, a rather advanced tool of probabilistic combinatorics.

In this short note, I give an elementary proof of the following, stronger result: writing V for the vertex set of the graph, for any source set S , the events $\{S \rightsquigarrow i\}$, $i \in V$, are positively associated – meaning that the expectation of the product of increasing functionals of the family $\{S \rightsquigarrow i\}$ for $i \in V$ is greater than the product of their expectations.

To show how this result can be used in concrete calculations, I also detail the example of percolation from the leaves of the randomly oriented complete binary tree of height n . Positive association makes it possible to use the Stein–Chen method to find conditions for the size of the percolation cluster to be Poissonian in the limit as n goes to infinity.

Contents

1	Introduction	2
1.1	Positive association and related notions	2
1.2	Notation	4
2	Positive association of the percolation cluster	4
2.1	Preliminary lemma	4
2.2	Main result	5
2.3	Corollaries	5
3	Percolation from the leaves of a binary tree	6
3.1	Setting and notation	6
3.1.1	The binary tree T_n	6
3.1.2	Percolation and downwards percolation on T_n	8
3.2	General results	9
3.2.1	Percolation threshold	9
3.2.2	Expected size of the percolation cluster	9
3.3	Badly-subcritical regime	11
3.3.1	Asymptotic cluster size and maximum depth	11
3.3.2	Second moments and main result	14

^{*}Center for Interdisciplinary Research in Biology (CIRB), CNRS UMR 7241, Collège de France, Paris, France.

[†]Laboratoire de Probabilités et Modèles Aléatoires (LPMA), CNRS UMR 7599, UPMC, Paris, France.

E-mail: francois.bienvenu@normalesup.org

1 Introduction

Oriented percolation is the study of connectivity in a random oriented graph. In most settings, one starts from a graph with a fixed orientation and then keeps each edge with a given probability. Classical such models include the north-east lattice [4] and the hypercube [6].

Another broad and natural class of random oriented graphs is obtained by starting from a fixed graph and then orienting each edge, independently of the orientations of other edges. Note that, in the general case, the orientations of the edges need not be unbiased: some edges can be allowed to have a higher probability to point towards one of their ends than towards the other. Percolation on such *randomly oriented graphs* has been studied, e.g. in [8], and more recently in [9], which motivated the present work.

In [9], Narayanan showed that if the edges of any fixed graph are randomly and independently oriented, then writing $\{S \rightsquigarrow i\}$ to indicate that there is an oriented path going from a vertex $s \in S$ to vertex i , we have

$$\mathbb{P}(S \rightsquigarrow i, S \rightsquigarrow j) \geq \mathbb{P}(S \rightsquigarrow i) \mathbb{P}(S \rightsquigarrow j).$$

The aim of this note is to strengthen and simplify the proof of this result. More specifically, let V be the vertex set of the graph. We prove that the events $\{S \rightsquigarrow i\}$, $i \in V$, are positively associated, without resorting to advanced results such as the Ahlswede–Daykin inequality [1].

To illustrate the usefulness of positive association, we finish the paper by detailing a simple but non-trivial model of percolation, where we let water percolate from the leaves of the randomly oriented complete binary tree of height n . The combinatorially simple structure of this graph makes it possible to bound the variance of the size of the percolation cluster. A classic application of the Stein–Chen method [11, 3] then provides us with a bound on the total variation distance between the size of the percolation cluster and a Poisson variable. This method is applicable to any graph where a bound on the variance of the size of the percolation cluster can be obtained.

1.1 Positive association and related notions

There are many ways to formalize the idea of a positive dependency between the random variables of a family $\mathbf{X} = (X_i)_{i \in I}$. A straightforward, weak one is to ask that these variables be pairwise positively correlated, i.e.

$$\forall i, j \in I, \quad \mathbb{E}(X_i X_j) \geq \mathbb{E}(X_i) \mathbb{E}(X_j).$$

A much stronger condition, due to [5], is known as positive association.

In the following definition and throughout the rest of this document, we use bold letters to denote vectors, as in $\mathbf{X} = (X_i)_{i \in I}$, and we write $\mathbf{X} \leq \mathbf{X}'$ to say that $X_i \leq X'_i$ for all i . Finally, a function $f: \mathbb{R}^I \rightarrow \mathbb{R}$ is said to be *increasing* when $\mathbf{X} \leq \mathbf{X}' \implies f(\mathbf{X}) \leq f(\mathbf{X}')$.

Definition 1.1. The random vector $\mathbf{X} = (X_i)_{i \in I}$ is said to be *positively associated* when, for all increasing functions f and g ,

$$\mathbb{E}(f(\mathbf{X})g(\mathbf{X})) \geq \mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{X}))$$

whenever these expectations exist.

In the rest of this document and without further mention, we only consider test functions f and g for which $\mathbb{E}(f(\mathbf{X}))$, $\mathbb{E}(g(\mathbf{X}))$ and $\mathbb{E}(f(\mathbf{X})g(\mathbf{X}))$ exist.

When there is no possible confusion, we sometimes omit to explicitly distinguish between an event and its indicator variable. Thus we say that the events A_i , $i \in I$, are positively associated when the random variables $\mathbb{1}_{A_i}$ are. Similarly, a random subset R of the fixed set I is assimilated to the vector

$$\mathbf{R} = \left(\mathbb{1}_{\{i \in R\}} \right)_{i \in I}$$

so that R is positively associated when the events $\{i \in R\}$, $i \in I$, are. Clearly, this is also equivalent to saying that for any increasing functions f and g from the power set of I to \mathbb{R} ,

$$\mathbb{E}(f(R)g(R)) \geq \mathbb{E}(f(R))\mathbb{E}(g(R)),$$

where f being increasing is understood to mean that $r' \subset r \implies f(r') \leq f(r)$.

Positive association is famous for the FKG theorem, which states that it is implied by a lattice condition that can sometimes be very easy to check [7]. In this document however, we will work directly from Definition 1.1.

Positive association has many applications, in part because – being a very strong form of positive dependence – it implies other more targeted positive dependence conditions. One such condition is positive relation.

Definition 1.2. The vector of Bernoulli variables $\mathbf{X} = (X_i)_{i \in I}$ is said to be *positively related* when, for each $i \in I$, there exists $\mathbf{X}^{(i)}$, built on the same space as \mathbf{X} , such that

- (i) $\mathbf{X}^{(i)}$ has the conditional distribution of \mathbf{X} given $\{X_i = 1\}$.
- (ii) $\mathbf{X}^{(i)} \geq \mathbf{X}$.

Equivalent characterizations of positive relation as well as a proof of the fact that it is implied by positive association can be found, e.g., in [10].

Positive relation is very useful when using the Stein–Chen method, as the next theorem shows. Proofs of this classic result and a more general introduction to the Stein–Chen method can be found in [2] and [10].

Theorem A (Stein–Chen method). *Let X_1, \dots, X_n be positively related Bernoulli variables, and $p_i = \mathbb{P}(X_i = 1)$. Let $W = \sum_{i=1}^n X_i$ and $\lambda = \mathbb{E}(W)$. Then,*

$$d_{\text{TV}}(W, \text{Poisson}(\lambda)) \leq \min\{1, \lambda^{-1}\} \left(\text{Var}(W) - \lambda + 2 \sum_{i=1}^n p_i^2 \right),$$

where d_{TV} denotes the total variation distance.

1.2 Notation

Let us fix some notation to be used throughout the rest of this document.

We study the simple graph $G = (V, E)$. Unless explicitly specified otherwise, V is assumed to be finite and we denote by $|V|$ its cardinal. The edges of G have a random orientation that is independent of the orientations of other edges and we write $\{i \rightarrow j\}$ to indicate that the edge $\{ij\}$ is oriented towards j . Formally, we are thus given a family of events $\{i \rightarrow j\}, \{ij\} \in E$, such that $\{i \rightarrow j\} = \{j \rightarrow i\}^c$ and for all $\{ij\} \neq \{kl\} \in E$, $\{i \rightarrow j\} \perp \{k \rightarrow \ell\}$.

Finally, for every vertices i and j , we write $\{i \rightsquigarrow j\}$ for the event that there exists an oriented path going from i to j . Similarly, for every source set S we let $\{S \rightsquigarrow i\} = \bigcup_{j \in S} \{j \rightsquigarrow i\}$ be the event that there is an oriented path from S to i , and for every target set T we let $\{i \rightsquigarrow T\} = \bigcup_{j \in T} \{i \rightsquigarrow j\}$ be the event that there is an oriented path from i to T . If there is an ambiguity regarding which graph is considered for these events, we will specify it with the notation $\{i \overset{G}{\rightsquigarrow} j\}$.

2 Positive association of the percolation cluster

2.1 Preliminary lemma

Lemma 2.1. *Let Γ be a finite set and let R be a positively associated random subset of Γ . Let $X_i^r, r \subset \Gamma$ and $i \in V$, be a family of events on the same probability space as R with the property that*

- (i) $r' \subset r \implies X_i^{r'} \subset X_i^r, \forall i \in V$.
- (ii) For all $r \subset \Gamma$, $(X_i^r)_{i \in V}$ is positively associated and independent of R .

For all $i \in V$, define X_i^R by

$$X_i^R := \bigcup_{r \subset \Gamma} \{R = r\} \cap X_i^r.$$

Then, the events $X_i^R, i \in V$, are positively associated.

Proof. Let f and g be two increasing functions. We have

$$\begin{aligned} \mathbb{E}(f(\mathbf{X}^R)g(\mathbf{X}^R)) &= \sum_{r \subset \Gamma} \mathbb{E}(f(\mathbf{X}^r)g(\mathbf{X}^r)\mathbb{1}_{\{R=r\}}) \\ &= \sum_{r \subset \Gamma} \mathbb{E}(f(\mathbf{X}^r)g(\mathbf{X}^r)) \mathbb{P}(R = r) \\ &\geq \sum_{r \subset \Gamma} \mathbb{E}(f(\mathbf{X}^r)) \mathbb{E}(g(\mathbf{X}^r)) \mathbb{P}(R = r), \end{aligned}$$

because $\mathbf{X}^r \perp R$ and \mathbf{X}^r is positively associated. Now, let $u: r \mapsto \mathbb{E}(f(\mathbf{X}^r))$ and $v: r \mapsto \mathbb{E}(g(\mathbf{X}^r))$, so that the last sum is $\mathbb{E}(u(R)v(R))$. Note that u and v are increasing, since f and g are and, by hypothesis, $r' \subset r \implies \mathbf{X}^{r'} \leq \mathbf{X}^r$. Therefore, by the positive association of R ,

$$\mathbb{E}(u(R)v(R)) \geq \mathbb{E}(u(R)) \mathbb{E}(v(R)).$$

Finally, using again the independence of \mathbf{X}^r and R , we have $\mathbb{E}(u(R)) = \mathbb{E}(f(\mathbf{X}^R))$ and $\mathbb{E}(v(R)) = \mathbb{E}(g(\mathbf{X}^R))$, which terminates the proof. \square

2.2 Main result

Theorem 2.2. *Let G be a finite graph with vertex set V , whose edges have been randomly and independently oriented. Then, for any source set S , the events $\{S \rightsquigarrow i\}$, $i \in V$, are positively associated, i.e., for all increasing functions f and g and writing $\mathbf{X} = (\mathbb{1}_{\{S \rightsquigarrow i\}})_{i \in V}$,*

$$\mathbb{E}(f(\mathbf{X})g(\mathbf{X})) \geq \mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{X})).$$

Proof. Our proof is essentially the same as Narayanan, i.e. we use the same induction on the number of vertices. The difference is that we use Lemma 2.1 rather than the Ahlswede–Daykin inequality to propagate the positive dependence.

The theorem is trivial for the graph consisting of a single vertex (a family of a single variable being positively associated) so let us assume that it holds for every graph with strictly less than $|V|$ vertices.

Let Γ be the neighborhood of S , i.e.

$$\Gamma = \left\{v \in V \setminus S : \exists s \in S \text{ s.t. } \{vs\} \in E\right\}.$$

Then, let R be the random subset of Γ defined by

$$R = \left\{v \in \Gamma : \exists s \in S \text{ s.t. } s \rightarrow v\right\}.$$

Observe that the events $\{i \in R\}$, $i \in \Gamma$ are independent, so that the set R is positively associated.

Next, let H be the subgraph of G induced by $V \setminus S$. Note that, for all $i \in V \setminus S$,

$$\{S \overset{G}{\rightsquigarrow} i\} = \{R \overset{H}{\rightsquigarrow} i\}.$$

For every fixed $r \subset \Gamma$, the family $\{r \overset{H}{\rightsquigarrow} i\}$ for $i \in V \setminus S$ is independent of R because it depends only on the orientations of the edges of H , while R depends only on the orientations of the edges of G that go from S to Γ – and these two sets of edges are disjoint. Moreover, by the induction hypothesis, the events $\{r \overset{H}{\rightsquigarrow} i\}$, $i \in V \setminus S$, are positively associated. Since for fixed sets r and r' such that $r' \subset r$, $\{r' \rightsquigarrow i\} \implies \{r \rightsquigarrow i\}$ for all vertices, we can apply Lemma 2.1 to conclude that the events $\{R \rightsquigarrow i\}$, $i \in V \setminus S$, are positively associated.

To terminate the proof, note that the events $\{S \rightsquigarrow i\}$ are certain for $i \in S$ and that the union of a family of positively associated events and of a family of certain events is still positively related. \square

2.3 Corollaries

Corollary 2.3. *Let G be a finite graph with independently oriented edges. For any target set T , the events $\{i \rightsquigarrow T\}$, $i \in V$, are positively associated.*

Proof. Consider the randomly oriented graph H obtained by reversing the orientation of the edges of G , i.e. such that $\{i \xrightarrow{H} j\} = \{j \xrightarrow{G} i\}$. Then for all $i \in V$,

$$\{i \xrightarrow{G} T\} = \{T \xrightarrow{H} i\},$$

and we already know from Theorem 2.2 that the events $\{T \xrightarrow{H} i\}$, $i \in V$, are positively associated. \square

Corollary 2.4. *Let G be an infinite graph with independently oriented edges. Let f and g be increasing, non-negative functions on \mathbb{R}^V that depend only on a finite number of coordinates (i.e. such that there exists a finite set $U \subset V$ and $\tilde{f}: \mathbb{R}^U \rightarrow [0, +\infty[$ such that $f = \tilde{f} \circ \varphi$, where φ is the canonical surjection from \mathbb{R}^V to \mathbb{R}^U). Then, for any source set S , letting $\mathbf{X} = (\mathbb{1}_{\{S \rightsquigarrow i\}})_{i \in V}$,*

$$\mathbb{E}(f(\mathbf{X})g(\mathbf{X})) \geq \mathbb{E}(f(\mathbf{X}))\mathbb{E}(g(\mathbf{X})).$$

Proof. Let G_n be an increasing sequence of graphs such that $G = \bigcup_n G_n$, and for all $i \in V$, let

$$X_i^{(n)} = \{S \xrightarrow{G_n} i\},$$

so that $X_i^{(n)} \subset X_i^{(n+1)}$ and $X_i = \bigcup_n X_i^{(n)}$. Since the functions f and g are increasing, so are the sequences $f(\mathbf{X}^{(n)})$ and $g(\mathbf{X}^{(n)})$. Thus, using Theorem 2.2 and monotone convergence,

$$\mathbb{E}\left(\lim_n f(\mathbf{X}^{(n)})g(\mathbf{X}^{(n)})\right) \geq \mathbb{E}\left(\lim_n f(\mathbf{X}^{(n)})\right)\mathbb{E}\left(\lim_n g(\mathbf{X}^{(n)})\right).$$

Finally, if f and g depend on a finite number of events X_i , then for every realization of \mathbf{X} we have $\lim_n f(\mathbf{X}^{(n)}) = f(\mathbf{X})$ and $\lim_n g(\mathbf{X}^{(n)}) = g(\mathbf{X})$. \square

Corollary 2.5 (Narayanan, 2016). *For any (possibly infinite) graph with independently oriented edges, for any source set S and for any two vertices i and j ,*

$$\mathbb{P}(S \rightsquigarrow i, S \rightsquigarrow j) \geq \mathbb{P}(S \rightsquigarrow i)\mathbb{P}(S \rightsquigarrow j)$$

Proof. Take $f: (x_k)_{k \in V} \mapsto x_i$ and $g: (x_k)_{k \in V} \mapsto x_j$ in Corollary 2.4. \square

3 Percolation from the leaves of a binary tree

In this section, we study percolation on the randomly oriented complete binary tree of height n . We start by introducing this graph and some notation.

3.1 Setting and notation

3.1.1 The binary tree T_n

Let V_n be the set of words of length at most n on the alphabet $\{0, 1\}$, i.e.

$$V_n = \bigcup_{k=0}^n \{0, 1\}^k,$$

where $\{0, 1\}^0$ is understood to represent the empty word.

A word v is said to be a *successor* of u when $v = us$, with $s \in \{0, 1\}$. Thus, every word of length less than n has two successors in V_n . Similarly, every non-empty word of V_n has exactly one predecessor. With this terminology, let

$$E_n = \{ \{u, v\} : (u, v) \in V_n^2, v \text{ is a successor of } u \}.$$

What we call the complete binary tree of height n is the graph T_n defined by $T_n = (V_n, E_n)$. Let us fix some vocabulary and notation for working with T_n .

The *leaves* of T_n are the vertices of degree 1, and its *root* is the only vertex of degree 2. The root will always be denoted by r .

The *level* of a vertex is its distance from the leaf set. Thus, the leaves are the level-0 vertices, and the root is the only vertex of level n . We will write $\ell(v)$ for the level of vertex v .

The unique path between two vertices u and v will be denoted by $[u, v]$. Sometimes, we will need to remove one of its ends from $[u, v]$, in which case we will write $]u, v[$ for $[u, v] \setminus \{u\}$ and $]u, v[$ for $[u, v] \setminus \{v\}$.

Finally, there is a natural order \preceq on the vertices of T_n , defined, e.g. by

$$u \preceq v \iff v \in [u, r]$$

Thus ordered, (V_n, \preceq) is a join-semilattice, i.e. we can define the *join* of any u and v , denoted by $u \vee v$, as

$$u \vee v = \inf ([u, r] \cap [v, r]) = \sup [u, v]$$

These definitions are illustrated in Figure 1A.

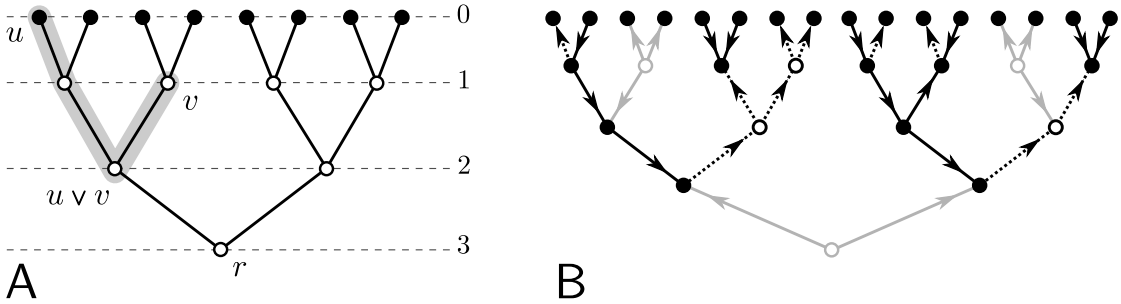


Figure 1: A, the complete binary tree T_3 . The black vertices are the leaves of the tree, and r is the root. The numbers on the right indicate the levels of the vertices. The path $[u, v]$ between u and v has been highlighted and $u \vee v$, the join of u and v , can be seen to be the unique vertex of maximum level in $[u, v]$. B, percolation and downwards percolation on T_4 . Water starts from the leaves and then flows downwards through black edges and upwards through dotted edges. It does not reach the grayed-out portions of the tree. The percolation cluster C_n consists of both black vertices and white vertices, while the downwards percolation cluster C_n^\downarrow consists of black vertices only. Note that the leaves are excluded from both percolation clusters.

3.1.2 Percolation and downwards percolation on T_n

Let every edge of T_n be oriented towards the root with probability p and towards the leaf set with probability $1 - p$, independently of the other edges.

In this application, the source set L will be the leaf set of T_n . In other words, we pump water into the leaves of T_n and let it flow through those edges whose orientation matches that of the flow, as depicted in Figure 1B. For any vertex v , write

$$X_v = \{L \rightsquigarrow v\},$$

for the event that the water reaches v , and

$$\pi_k^{(n)} = \mathbb{P}(X_v), \quad \text{where } k = \ell(v)$$

for the probability of this event. In the special case where $v = r$ is the root, we use the notation

$$\rho_n = \pi_n^{(n)} = \mathbb{P}(X_r).$$

Finally, let

$$C_n = \{v \in V_n \setminus L : X_v\}$$

denote the percolation cluster.

As will become clear, this percolation model is closely related to a simpler one where, in addition to respecting the orientation of edges, water is constrained to flow towards increasing levels of the tree. If we think of the root of T_n as its bottom and of the leaves as its top, then water runs down from the leaves, traveling through downwards-oriented edges; hence we refer to this second model as *downwards percolation*. Again, this is represented in Figure 1B.

Let us write Y_v for the event that that vertex v gets wet in downwards percolation, and let

$$C_n^\downarrow = \{v \in V_n \setminus L : Y_v\}$$

be the downwards-percolation cluster.

How are percolation and downwards percolation related? First, it follows directly from the definition that $Y_v \subset X_v$. Second, note that

$$Y_r = X_r$$

because every path from the leaf set to r is downwards-oriented. Furthermore, letting $T_n^{(\preceq v)}$ denote the subtree of T_n induced by v and the vertices that are above it, then the randomly oriented trees $T_n^{(\preceq v)}$ and $T_{\ell(v)}$ have the same law. As a result, for all $v \in V_n$,

$$\mathbb{P}(Y_v) = \rho_{\ell(v)},$$

from which the next proposition follows.

Proposition 3.1. *Let $|C_n^\downarrow|$ be the number of wet vertices (not counting the leaves) in the downwards-percolation model on T_n . We have*

$$\mathbb{E}\left(|C_n^\downarrow|\right) = \sum_{k=1}^n 2^{n-k} \rho_k.$$

3.2 General results

3.2.1 Percolation threshold

If the probability p that an edge is oriented towards the root is sufficiently small, the probability ρ_n of the root getting wet will go to zero as n goes to infinity. Define the percolation threshold as

$$\theta_c = \sup \{p : \rho_n \rightarrow 0\}.$$

Proposition 3.2. *The probability ρ_n of the root of T_n getting wet in either percolation model satisfies the following recurrence:*

$$\rho_{n+1} = 2p\rho_n - (p\rho_n)^2, \quad \text{with } \rho_0 = 1.$$

The percolation threshold is therefore $\theta_c = 1/2$ and

- (i) $p \leq \theta_c \implies \rho_n \rightarrow 0$.
- (ii) $p > \theta_c \implies \rho_n \rightarrow (2p - 1)/p^2 > 0$.

Proof. First, note that

$$Y_r = (Y_0 \cap \{0 \rightarrow r\}) \cup (Y_1 \cap \{1 \rightarrow r\}),$$

where 1 and 0 are the two successors of the root r . These four events are independent and we have $\mathbb{P}(0 \rightarrow r) = \mathbb{P}(1 \rightarrow r) = p$ and $\mathbb{P}(Y_0) = \mathbb{P}(Y_1) = \rho_{n-1}$, whence the recurrence relation.

Now, let $f_p: x \mapsto 2px - (px)^2$, so that $\rho_{n+1} = f_p(\rho_n)$. For $p \leq 1/2$, the only solution to the equation $f_p(x) = x$ in $[0, 1]$ is $x = 0$, and $f_p(x) < x$ for all $0 < x \leq 1$. This proves (i). For $p > 1/2$, the equation $f_p(x) = x$ has a non-zero solution $\alpha = (2p - 1)/p^2$ in $[0, 1]$. Finally, $f_p(x) > x$ for $0 < x < \alpha$ and $f_p(x) < x$ for $\alpha < x < 1$, proving (ii). \square

Remark. Another way to obtain Proposition 3.2 is to note that the existence of an open path from the leaf set of T_n to its root is equivalent to the existence of a path of length n starting from the root of a Galton–Watson tree with Binomial(2, p) offspring distribution, i.e. to its non-extinction after n generations. In the limit as $n \rightarrow \infty$, the probability of non-extinction is strictly positive if and only if the expected number of offspring is greater than 1 – i.e., in our case, $2p > 1$.

3.2.2 Expected size of the percolation cluster

Let us clarify the relation between percolation and downwards percolation by expressing the probability $\pi_k^{(n)}$ that a vertex gets wet in (bidirectional) percolation as a function of ρ_k, \dots, ρ_n .

Proposition 3.3. *Let $\pi_k^{(n)} = \mathbb{P}(X_v)$, where $\ell(v) = k$. We have*

$$\pi_k^{(n)} = \rho_k + (1 - \rho_k) \alpha_k^{(n)},$$

where

$$\alpha_k^{(n)} = (1-p)p \sum_{i=0}^{n-1-k} (1-p)^i \rho_{k+i} \prod_{j=0}^{i-1} (1-p\rho_{k+j})$$

is the probability that water reaches v “from below” and ρ_k is the probability that it reaches it “from above”.

Remark. To make sense of the expression of $\alpha_k^{(n)}$, it can also be written as

$$\alpha_k^{(n)} = \sum_{i=1}^{n-k} \mathbb{P}(M = k+i), \quad \text{with } \mathbb{P}(M = k+i) = (1-p)^i p \rho_{k+i-1} \prod_{j=0}^{i-2} (1-p\rho_{k+j}).$$

In this expression, M is the level of the highest (that is, minimal with respect to \preceq) vertex $u \in]v, r]$ such that $Y_u \cap \{u \rightsquigarrow v\}$ (with $M = +\infty$ if there is no such vertex).

Proof. Water can reach v from above (i.e. coming from one of its successors) or from below (coming from its predecessor). These two events are independent, because they depend on what happens in disjoint regions of T_n .

Water reaches v from above if and only if v gets wet in downwards percolation. To reach v from below, water had to travel through a portion of the path $[v, r]$ from v to the root. To enter this portion of the path, it had to reach at least one vertex, say u , from above. Let $\varphi(u)$ be the successor of u that does not belong to $[v, r]$. The water had to get to $\varphi(u)$ from above, flow to u , and from here to v .

This reasoning, which is illustrated in Figure 2A, leads us to rewrite X_v as

$$X_v = Y_v \cup \bigcup_{u \in]v, r]} \left(Y_{\varphi(u)} \cap \{\varphi(u) \rightarrow u\} \cap \{u \rightsquigarrow v\} \right)$$

In order to compute the probability of this event, we rewrite it as the disjoint union

$$X_v = Y_v \cup \bigcup_{u \in]v, r]} \left(Y_v^c \cap \left(\bigcap_{w \in]v, u[} \tilde{Y}_w^c \right) \cap \tilde{Y}_u \cap \{u \rightsquigarrow v\} \right),$$

where

$$\tilde{Y}_x = Y_{\varphi(x)} \cap \{\varphi(x) \rightarrow x\}.$$

Next, we note that the factors of each term of the union over $u \in]v, r]$ are independent, because they are determined by the orientations of disjoint sets of edges: Y_v depends only on the orientations of the edges of $T_n^{(\preceq v)}$; each \tilde{Y}_x of those of the edges of $T_n^{(\preceq \varphi(x))}$ and of $\{x, \varphi(x)\}$; and $\{u \rightsquigarrow v\}$ of the edges of $[u, v]$. Using that $\mathbb{P}(Y_v) = \rho_{\ell(v)}$, $\mathbb{P}(\tilde{Y}_x) = p\rho_{\ell(x)-1}$ and $\mathbb{P}(u \rightsquigarrow v) = (1-p)^{d(u,v)}$ and replacing the sum on the vertices of $]v, r]$ by a sum on their levels, we get the desired expression. \square

From Proposition 3.3, we get the following expression for the expected size of the percolation cluster:

Proposition 3.4. *Let $|C_n|$ be the number of wet vertices, not counting the leaves, in the (bidirectional) percolation model on T_n . Then,*

$$\mathbb{E}(|C_n|) = \sum_{k=1}^n 2^{n-k} \left(\rho_k + (1-\rho_k) \alpha_k^{(n)} \right),$$

where $\alpha_k^{(n)}$ is defined in Proposition 3.3.

Using a similar reasoning, it is also possible to express $\mathbb{P}(X_u X_v)$ – and from there $\text{Var}(|C_n|)$ – as a function of p and ρ_1, \dots, ρ_n only. However, the resulting expression is rather complicated, and thus of little interest. We will therefore only give the asymptotic estimates that are needed to apply Theorem 2.2.

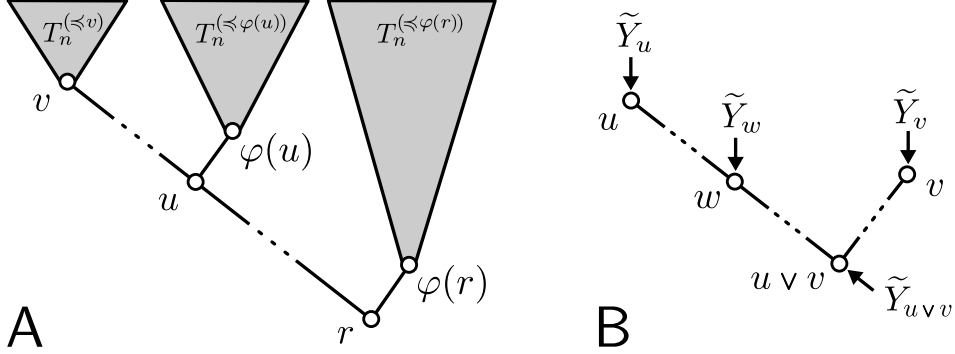


Figure 2: A, the notations used in the proof of Proposition 3.3. Water can reach v from above, i.e. traveling through $T_n^{(\leq v)}$, or from below, coming from some vertex $u \in]v, r]$. B, the notations used in the proof of Proposition 3.6. The arrows represent possible entry points for the water, and the \tilde{Y}_x the associated events, i.e., \tilde{Y}_x is the event that x receives water from the corresponding arrow.

3.3 Badly-subcritical regime

In this section, we focus on what happens when $p = p_n$ is allowed to depend on n and made to go to zero as n goes to infinity. We are therefore in a “badly-subcritical” regime, where only a negligible fraction of the vertices are going to get wet.

Note that the results of the previous sections still hold, provided that ρ_k is understood to depend on n as the solution of

$$\rho_{k+1} = 2p_n \rho_k - (p_n \rho_k)^2, \quad \rho_0 = 1.$$

To avoid clutter, the dependence in n will remain implicit and we will keep the notation ρ_k .

3.3.1 Asymptotic cluster size and maximum depth

Proposition 3.5. *When $p_n \rightarrow 0$, then as $n \rightarrow \infty$,*

$$\rho_k \sim (2p_n)^k,$$

where the convergence is uniform in k .

Proof. Clearly,

$$\rho_k \leq (2p_n)^k.$$

Plugging this first inequality into the recurrence relation for ρ_k , we get

$$\rho_{k+1} \geq (2p_n - p_n^2(2p_n)^k)\rho_k,$$

from which it follows that

$$\begin{aligned} \rho_k &\geq (2p_n)^k \prod_{i=0}^{k-1} \left(1 - \frac{p_n}{2}(2p_n)^i\right) \\ &\geq (2p_n)^k \prod_{i=1}^k \left(1 - (2p_n)^i\right), \end{aligned}$$

Let us show that

$$P_n^{(k)} = \prod_{i=1}^k \left(1 - (2p_n)^i\right)$$

has a lower bound that goes to 1 uniformly in k as $n \rightarrow \infty$. For all $k \geq 1$,

$$\log(P_n^{(k)}) \geq \sum_{i=1}^{\infty} \log(1 - (2p_n)^i)$$

Now,

$$\sum_{i=1}^{\infty} \log(1 - (2p_n)^i) = -\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \frac{1}{j} (2p_n)^{ij} \geq -\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (2p_n)^{ij}$$

and

$$-\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (2p_n)^{ij} = -\sum_{i=1}^{\infty} \frac{(2p_n)^i}{1 - (2p_n)^i} \geq -\sum_{i=1}^{\infty} \frac{(2p_n)^i}{1 - 2p_n} = -\frac{2p_n}{(1 - 2p_n)^2},$$

so that $P_n^{(k)} \geq \exp(-2p_n/(1 - 2p_n)^2)$. Putting the pieces together,

$$e^{-\frac{2p_n}{(1-2p_n)^2}} (2p_n)^k \leq \rho_k \leq (2p_n)^k,$$

which terminates the proof. \square

Proposition 3.6. *When $p_n \rightarrow 0$, then as $n \rightarrow \infty$,*

$$\mathbb{E}(|C_n|) \sim \mathbb{E}(|C_n^\downarrow|) \sim 2^n p_n.$$

Proof. From the expression of $\alpha_k^{(n)}$ given in Proposition 3.3,

$$\alpha_k^{(n)} \leq p_n \sum_{i=0}^{n-1-k} \rho_{k+i}.$$

But since $\rho_{k+i} \leq (2p_n)^i \rho_k$,

$$\alpha_k^{(n)} \leq \frac{p_n \rho_k}{1 - 2p_n}.$$

Using this in Propositions 3.1 and 3.4, we see that for n large enough,

$$\mathbb{E}(|C_n|) \leq \left(1 + \frac{p_n}{1 - 2p_n}\right) \mathbb{E}(|C_n^\downarrow|),$$

Next, using again that $\rho_k \leq (2p_n)^{k-1} \rho_1$,

$$2^{n-1} \rho_1 \leq \sum_{k=1}^n 2^{n-k} \rho_k \leq 2^{n-1} \rho_1 \sum_{k=1}^n p_n^{k-1}.$$

Since the sum in right-hand side is bounded above by $1/(1-p_n)$ and since $\rho_1 \sim 2p_n$, this finishes the proof. \square

Proposition 3.6 shows that, in the badly-subcritical regime, the overwhelming majority of wet vertices are level-1 vertices. It is therefore natural to wonder: how deep does water go?

Proposition 3.7. *Let $\ell_{\max}^{(n)}$ be the maximum level reached by water, and let*

$$\kappa_n = \frac{\log(2)n}{\log(1/p_n)}.$$

If $p_n \rightarrow 0$, then letting $\lfloor x \rfloor = \lfloor x + 1/2 \rfloor$ denote the nearest integer to x ,

$$\mathbb{P}\left(\ell_{\max}^{(n)} = \lfloor \kappa_n \rfloor - 1 \text{ or } \lfloor \kappa_n \rfloor\right) \rightarrow 1$$

as $n \rightarrow \infty$. In particular,

- *If $p_n = n^{-\alpha}$, then $\kappa_n = cn/\log(n)$, with $c = \log(2)/\alpha$.*
- *If $p_n = \gamma^{-n}$, $1 < \gamma \leq 2$, then $\kappa_n = \frac{\log(2)}{\log(\gamma)}$.*

Proposition 3.7 shows that the maximum level reached by water is remarkably deterministic in the limit as n goes to infinity, independently of the speed of convergence of p_n to zero. It also shows that, even in the badly-subcritical regime, water can go infinitely deep – even though these depths will always represent a negligible fraction of the total height of T_n .

Before jumping to the proof, let us give a simple heuristic. Let

$$B_k^{(n)} = \text{Card}\{v \in C_n^\downarrow : \ell(v) = k\}$$

be the number of level- k vertices that get wet in downwards-percolation. Using Proposition 3.5, we see that

$$\mathbb{E}(B_k^{(n)}) = \rho_k 2^{n-k} \sim (p_n)^k 2^n.$$

If k is such that this expectation goes to zero, then the probability that this level will be reached by water will go to zero and k will be a lower bound on $\ell_{\max}^{(n)}$. Conversely, if this expectation goes to infinity then it seems reasonable to expect that $B_k^{(n)} \geq 1$ with high probability, in which case we would have $\ell_{\max}^{(n)} \geq k$.

Proof. Let L_k be the set of level- k vertices. The event that water does not reach level k is

$$\{\ell_{\max}^{(n)} < k\} = \bigcap_{v \in L_k} Y_v^c.$$

Since each Y_v depends only on $T_n^{(\preceq v)}$, these events are independent and we have

$$\mathbb{P}(\ell_{\max}^{(n)} < k) = (1 - \rho_k)^{2^{n-k}}.$$

Whether this expression goes to 0 or to 1 is determined by whether $\rho_k 2^{n-k}$ goes to $+\infty$ or to 0, respectively. Now let $k = k_n$ depend on n . By Proposition 3.5, we have

$$\rho_{k_n} 2^{n-k_n} \sim (p_n)^{k_n} 2^n.$$

Again, whether this quantity goes to $+\infty$ or to 0 depends on whether

$$w_n = \log(p_n) k_n + \log(2) n$$

goes to $+\infty$ or to $-\infty$, respectively. Setting $\kappa_n = \frac{\log(2)n}{\log(1/p_n)}$, we see that:

- (i) If there exists $\eta > 0$ such that $k_n < \kappa_n - \eta$ for all n , then $w_n \rightarrow +\infty$ and as a result $\mathbb{P}(\ell_{\max}^{(n)} \geq k) \rightarrow 1$.
- (ii) If there exists $\eta > 0$ such that $k_n > \kappa_n + \eta$ for all n , then $w_n \rightarrow -\infty$ and as a result $\mathbb{P}(\ell_{\max}^{(n)} < k) \rightarrow 1$.

Finally, we note that

- $\lfloor \kappa_n \rfloor - 1 < \kappa_n - 1/2$. By (i), this shows that $\mathbb{P}(\ell_{\max}^{(n)} \geq \lfloor \kappa_n \rfloor - 1) \rightarrow 1$.
- $\lfloor \kappa_n \rfloor + 1 \geq \kappa_n + 1/2$. By (ii), this shows that $\mathbb{P}(\ell_{\max}^{(n)} < \lfloor \kappa_n \rfloor + 1) \rightarrow 1$.

As a result,

$$\mathbb{P}(\lfloor \kappa_n \rfloor - 1 \leq \ell_{\max}^{(n)} \leq \lfloor \kappa_n \rfloor) \rightarrow 1,$$

and the proof is complete. \square

3.3.2 Second moments and main result

We have seen in the previous section that level-1 vertices account for a fraction 1 of the expected size of both percolation clusters. But do they also account for a fraction 1 of the variances?

For downwards percolation, it is not hard to convince oneself that it is. Indeed, the number $B_1^{(n)}$ of wet vertices of level 1 is a binomial variable with parameters ρ_1 and 2^{n-1} . From here, if we neglect “collision” events, where a vertex receives water from both vertices immediately above it, then the downwards percolation cluster resembles $B_1^{(n)}$ independent paths with geometric lengths, that is,

$$|C_n^\downarrow| \approx \sum_{i=1}^{B_1^{(n)}} \tau_i, \quad \text{where } \tau_i \sim \text{Geometric}(1 - p_n).$$

Since $\text{Var}(\tau_i) \sim p_n$, by a simple application of the law of total variance we find that

$$\text{Var}(|C_n^\downarrow|) \approx \text{Var}(B_1^{(n)}) \sim p_n 2^n.$$

For bidirectional percolation however, things are not so obvious because there is a very strong feedback from higher-level vertices to lower-level ones: if vertex v gets wet, water will flow up from it to most vertices of $T_n^{(\preceq v)}$ that are not already wet. Thus, every rare event where water reaches a vertex of level k will result in approximately 2^k additional vertices getting wet – which it seems could increase the variance of $|C_n|$. However we will see that this is not the case.

Proposition 3.8. *In the regime $p_n \rightarrow 0$,*

- (i) *If $u \preceq v$ then $\text{Cov}(X_u, X_v) \leq \pi_{\ell(v)}^{(n)}$.*
- (ii) *Otherwise, $\text{Cov}(X_u, X_v) \leq d(u, v) \rho_{\ell(u \vee v)}$.*

As a result, $\text{Var}(|C_n|) \sim 2^n p_n$ and

$$\text{Var}(|C_n|) - \mathbb{E}(|C_n|) = O(2^n p_n^2).$$

Proof. Point (i) is clear, since

$$\text{Cov}(X_u, X_v) \leq \mathbb{P}(X_u \cap X_v) \leq \min\{\mathbb{P}(X_u), \mathbb{P}(X_v)\} = \pi_{\ell(v)}^{(n)}.$$

As a side note, this upper bound on $\mathbb{P}(X_u \cap X_v)$ is not as crude as it may seem. Indeed, $X_v \cap \{v \rightsquigarrow u\} \subset X_u \cap X_v$ and it is not hard to check that $\mathbb{P}(X_v \cap \{v \rightsquigarrow u\})$ is greater than $(1 - p_n)^{d(u,v)} \rho_{\ell(v)}/2 \sim \pi_{\ell(v)}^{(n)}/2$.

To prove (ii), let us show that

$$\mathbb{P}(X_u \cap X_v) \leq \pi_u^{(n)} \pi_v^{(n)} + d(u, v) \rho_{\ell(u \vee v)}.$$

As in the proof of Proposition 3.3, we start by re-expressing X_u . For every $w \in]u, v[$, let $\varphi(w)$ be the successor of w that does not belong to $[u, v]$, and for every $z \in]u \vee v, r]$, let $\psi(z)$ be the successor of z that does not belong to $[u \vee v, r]$. Then, for every $w \in [u, v]$, define \tilde{Y}_w by

$$\tilde{Y}_w = \begin{cases} Y_w & \text{if } w = u \text{ or } w = v. \\ \bigcup_{z \in]u \vee v, r]} Y_{\psi(z)} \cap \{\psi(z) \rightarrow z\} \cap \{z \rightsquigarrow u \vee v\} & \text{if } w = u \vee v. \\ Y_{\varphi(w)} \cap \{\varphi(w) \rightarrow w\} & \text{otherwise.} \end{cases}$$

These definitions are illustrated in Figure 2B. Note that $\tilde{Y}_{u \vee v}$ is simply the event that $u \vee v$ receives water “from below”. Thus, using the notation of Proposition 3.3, we have

$$\mathbb{P}(\tilde{Y}_{u \vee v}) = \alpha_{\ell(u \vee v)}^{(n)}.$$

For both $s = u$ and $s = v$ we have

$$X_s = \bigcup_{w \in [u, v]} \tilde{Y}_w \cap \{w \rightsquigarrow s\}.$$

Again, we rewrite this as the disjoint union

$$X_s = \bigcup_{w \in [u, v]} Z_w^s,$$

where

$$Z_w^s = \left(\bigcap_{z \in [s, w[} \tilde{Y}_z^c \right) \cap \tilde{Y}_w \cap \{w \rightsquigarrow s\}.$$

Next, we note that for any vertices x and y in $[u, v]$,

- If $[u, x] \cap [y, v] = \emptyset$, then $Z_x^u \perp Z_y^v$.
- If $[u, x] \cap [y, v] = \{w\}$, then $Z_x^u \cap Z_y^v = \tilde{Y}_w \cap \{w \rightsquigarrow u\} \cap \{w \rightsquigarrow v\}$.
- Otherwise, $Z_x^u \cap Z_y^v = \emptyset$.

As a result,

$$\begin{aligned} X_u \cap X_v &= \bigcup_{x \in [u, v]} \bigcup_{y \in [u, v]} Z_x^u \cap Z_y^v \\ &= \bigcup_{x \in [u, v]} \left(\left(\bigcup_{y \in [x, v]} Z_x^u \cap Z_y^v \right) \cup (Z_x^u \cap Z_x^v) \right) \end{aligned}$$

It follows that

$$\mathbb{P}(X_u, X_v) = \sum_{x \in [u, v]} \sum_{y \in [x, v]} \mathbb{P}(Z_x^u) \mathbb{P}(Z_y^v) + \mathbb{P}(\tilde{Y}_x, x \rightsquigarrow u, x \rightsquigarrow v).$$

To bound this sum, first note that

$$\sum_{x \in [u, v]} \sum_{y \in [x, v]} \mathbb{P}(Z_x^u) \mathbb{P}(Z_y^v) \leq \sum_{x \in [u, v]} \sum_{y \in [u, v]} \mathbb{P}(Z_x^u) \mathbb{P}(Z_y^v) = \pi_u^{(n)} \pi_v^{(n)}.$$

Next, \tilde{Y}_x and $\{x \rightsquigarrow u, x \rightsquigarrow v\}$ are independent and, writing $m(x) = d(x, u \vee v)$ for the number of downwards-oriented edges in the unique configuration of the edges of $[u, v]$ such that $\{x \rightsquigarrow u, x \rightsquigarrow v\}$,

$$\mathbb{P}(x \rightsquigarrow u, x \rightsquigarrow v) = p_n^{m(x)} (1 - p_n)^{d(u, v) - m(x)}$$

while

$$\mathbb{P}(\tilde{Y}_x) = \begin{cases} \alpha_{\ell(u \vee v)}^{(n)} & \text{if } x = u \vee v \\ \rho_{\ell(x)} & \text{if } x = u \text{ or } x = v \\ p_n \rho_{\ell(x) - 1} & \text{otherwise.} \end{cases}$$

Since for $x \in [u, v]$, $\ell(x) = \ell(u \vee v) - m(x)$, and that

$$\rho_{\ell(u \vee v) - k} \sim (2p_n)^{-k} \rho_{\ell(u \vee v)} \leq p_n^{-k} \rho_{\ell(u \vee v)},$$

we see that for every $x \in [u, v]$, $x \neq u \vee v$,

$$\mathbb{P}(\tilde{Y}_x, x \rightsquigarrow u, x \rightsquigarrow v) \leq \rho_{\ell(u \vee v)},$$

while for $x = u \vee v$ we already know from Proposition 3.6 and its proof that

$$\alpha_{\ell(u \vee v)}^{(n)} \leq p_n \rho_{\ell(u \vee v)}.$$

Discarding this negligible last contribution and summing these inequalities over the $d(u, v)$ vertices of $[u, v] \setminus \{u \vee v\}$, we find that

$$\sum_{x \in [u, v]} \mathbb{P}(\tilde{Y}_x, x \rightsquigarrow u, x \rightsquigarrow v) \leq d(u, v) \rho_{\ell(u \vee v)},$$

which complete the proof of (ii).

Now let us show that $\text{Var}(|C_n|) \leq \mathbb{E}(|C_n|) + O(2^n p_n^2)$. For $w \in]v, r]$, let $\varphi(w)$ denote the successor of w that does not belong to $[v, r]$. We decompose $\text{Var}(|C_n|)$ into

$$\text{Var}(|C_n|) = \sum_{v \in T_n} \left(\sum_{u \in T_n^{(\preceq v)}} \text{Cov}(X_v, X_u) + \sum_{w \in]v, r]} \left(\text{Cov}(X_v, X_w) + \sum_{x \in T_n^{(\preceq \varphi(w))}} \text{Cov}(X_v, X_x) \right) \right)$$

where it is understood that the sums exclude leaves. Using (i), we see that

$$\sum_{u \in T_n^{(\preceq v)}} \text{Cov}(X_v, X_u) \leq (2^{\ell(v)} - 1) \pi_{\ell(v)}^{(n)}.$$

Similarly, using (ii) we have

$$\begin{aligned} \sum_{x \in T_n^{(\leq \varphi(w))}} \text{Cov}(X_v, X_x) &\leq \sum_{x \in T_n^{(\leq \varphi(w))}} (d(v, w) + d(w, x)) \rho_{\ell(w)} \\ &\leq (2^{\ell(w)-1} - 1) (d(v, w) + \ell(w) - 1) \rho_{\ell(w)}. \end{aligned}$$

Since $\text{Cov}(X_v, X_w) \leq \pi_{\ell(w)}^{(n)}$, which is asymptotically equivalent to $\rho_{\ell(w)}$, we have

$$\text{Cov}(X_v, X_w) + \sum_{x \in T_n^{(\leq \varphi(w))}} \text{Cov}(X_v, X_x) \leq (\ell(w) + d(v, w)) 2^{\ell(w)} \rho_{\ell(w)}$$

Replacing the sum on w by a sum on its level and letting k denote the level of v , we get

$$\begin{aligned} \sum_{w \in]v, r]} \left(\text{Cov}(X_v, X_w) + \sum_{x \in T_n^{(\leq \varphi(w))}} \text{Cov}(X_v, X_x) \right) &\leq \sum_{i=1}^{n-k} (k + 2i) 2^{k+i} \rho_{k+i} \\ &\leq 2^k \rho_k \left(k \sum_{i=1}^{n-k} (4p_n)^i + 2 \sum_{i=1}^{n-k} i (4p_n)^i \right) \\ &\leq (1 + \varepsilon) p_n (k + 2) 2^{k+2} \rho_k \end{aligned}$$

For every $\varepsilon > 0$. Putting the pieces together, we find that

$$\text{Var}(|C_n|) \leq \sum_{k=1}^n 2^{n-k} (2^k - 1) \pi_k^{(n)} + (1 + \varepsilon) p_n \sum_{k=1}^n 2^{n+2} (k + 2) \rho_k.$$

The first sum is

$$\sum_{k=1}^n 2^{n-k} (2^k - 1) \pi_k^{(n)} = \mathbb{E}(|C_n|) + \sum_{k=1}^n 2^{n-(k-1)} (2^{k-1} - 1) \pi_k^{(n)}$$

where

$$\sum_{k=1}^n 2^{n-(k-1)} (2^{k-1} - 1) \pi_k^{(n)} \leq 2^n \sum_{k=2}^n \pi_k^{(n)} = O(2^n p_n^2),$$

since $\pi_k^{(n)} \leq (1 + \varepsilon) \rho_k$ and $\rho_k \leq (2p_n)^k$. Finally, the second sum is also clearly $O(2^n p_n^2)$, and the proof is complete. \square

With Proposition 3.8, Theorem 2.2 makes the following result immediate.

Proposition 3.9. *In the regime $p_n \rightarrow 0$, we have*

$$d_{\text{TV}}(|C_n|, \text{Poisson}(2^n p_n)) = O(p_n)$$

where d_{TV} denotes the total variation distance.

Proof. The proposition is a direct application of the Stein–Chen method (Theorem A) to the positively related variables X_v , $v \in T_n$. \square

Acknowledgments

I thank my advisor Amaury Lambert for a careful re-reading of the first draft of this manuscript. I also thank Jean-Jil Duchamps for helpful discussions.

References

- [1] R. Ahlswede and D. E. Daykin. An inequality for the weights of two families of sets, their unions and intersections. *Probability Theory and Related Fields*, 43(3):183–185, 1978.
- [2] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*. Oxford Studies in Probability. Oxford University Press, 1992.
- [3] L. H. Y. Chen. Poisson approximation for dependent trials. *Annals of Probability*, 3(3):534–545, 1975.
- [4] R. Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12(4):999–1040, 1984.
- [5] J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.
- [6] J. A. Fill and R. Pemantle. Percolation, first-passage percolation and covering times for Richardson’s model on the n-cube. *The Annals of Applied Probability*, 3(2):593–629, 1993.
- [7] C. M. Fortuin, P. W. Kasteleyn, and J. Ginibre. Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics*, 22(2):89–103, 1971.
- [8] S. Linusson. A note on correlations in randomly oriented graphs. *arXiv preprint arXiv:0905.2881*, 2009.
- [9] B. Narayanan. Connections in randomly oriented graphs. *Combinatorics, Probability and Computing*, pages 1–5, 2016.
- [10] N. Ross. Fundamentals of Stein’s method. *Probability Surveys*, 8:201–293, 2011.
- [11] C. M. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 2, pages 583–602. University of California Press, 1972.