Combinatorics, Probability and Computing (2011) 20, 1–9. © Cambridge University Press 2010
doi:10.1017/S0963548310000131

A Counter-Intuitive Correlation in a Random Tournament

SVEN ERICK ALM[1] and SVANTE LINUSSON[2†‡]

[1] Department of Mathematics, Uppsala University, PO Box 480, SE-751 06, Uppsala, Sweden
(e-mail: sea@math.uu.se)
[2] Department of Mathematics, KTH – Royal Institute of Technology, SE-100 44, Stockholm, Sweden
(e-mail: linusson@math.kth.se)

Received 4 September 2009; revised 3 March 2010; first published online 13 May 2010

Consider a randomly oriented graph $G = (V, E)$ and let a, s and b be three distinct vertices in V. We study the correlation between the events $\{a \to s\}$ and $\{s \to b\}$. We show that, counter-intuitively, when G is the complete graph K_n, $n \geqslant 5$, then the correlation is positive. (It is negative for $n = 3$ and zero for $n = 4$.) We briefly discuss and pose problems for the same question on other graphs.

1. Introduction

Given a graph $G = (V, E)$, we orient each edge with equal probability for the two possible directions and independently of all other edges. This model has been studied previously in, for instance, [6], [8] and [10]. Let a, s and b be three distinct vertices in V. The object of this paper is to study the correlation of the two events $\{a \to s\}$, that there exists a directed path from a to s, and $\{s \to b\}$. One might intuitively guess that they are always negatively correlated, *i.e.*, that $\mathbb{P}(a \to s, s \to b) < \mathbb{P}(a \to s) \cdot \mathbb{P}(s \to b)$. This is, however, not true in general.

It turns out to be easier to study the complementary events, $A := \{a \nrightarrow s\}$, that there does not exist a directed path from a to s, and $B := \{s \nrightarrow b\}$, which have the same covariance. In Theorem 2.1 we prove that for K_n, n large, the two dominating obstructions for a path from a to s are the event that all edges incident to a are directed towards a and the event that all edges incident to s are directed away from s. This fact might even strengthen one's belief that the events A and B are negatively correlated. However, in Section 2 we prove that for the complete graph, K_n, the events are positively correlated for $n \geqslant 5$, and we show that $\mathbb{P}(B|A)/\mathbb{P}(B)$ converges to $3/2$ as $n \to \infty$.

† Svante Linusson is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.
‡ This research was conducted when both authors visited the Institut Mittag-Leffler (Djursholm, Sweden).

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

In Section 3 we give exact recursions for the probabilities $\mathbb{P}(A)$ and $\mathbb{P}(A \cap B)$. For completeness, in Section 4 we show that the events are negatively correlated when G is a cycle and negatively correlated or independent when G is a tree, and give an example of a graph where the correlation is negative or positive depending on the choice of a, b, s. We end by stating a number of conjectures and open problems.

In a forthcoming paper, [2], we study this problem when G is the random graph $G(n, p)$.

Having proved something that at first surprised us, we have tried with hindsight to formulate a new heuristic understanding. One way to understand this result is that the event $a \nrightarrow s$ in K_n really has only two cases, of equal probability, that we need to consider. In the first case the outdegree of a is zero, which means that we can restrict to the graph on $[n]\backslash\{a\}$, where the event $s \nrightarrow b$ is twice as likely, heuristically speaking: see Theorem 2.1. In the second case the indegree of s is zero, which makes the event $s \nrightarrow b$ less likely (50% less in fact, since only the case of indegree zero for b remains); but it is still larger than zero, so together it gives an increase.

The question studied here was posed in [8]. There it was proved that, under this model, for any vertices $a, b, s, t \in V$ the events $\{s \to a\}$ and $\{s \to b\}$ are never negatively correlated. This was also shown to be true if we first condition on $\{s \nrightarrow t\}$, *i.e.*,

$$\mathbb{P}(s \to a, s \to b | s \nrightarrow t) \geqslant \mathbb{P}(s \to a | s \nrightarrow t) \cdot \mathbb{P}(s \to b | s \nrightarrow t).$$

As a sort of converse it was also proved that

$$\mathbb{P}(s \to a, b \to t | s \nrightarrow t) \leqslant \mathbb{P}(s \to a | s \nrightarrow t) \cdot \mathbb{P}(b \to t | s \nrightarrow t).$$

The proofs in [8] relied heavily on the results in [3] and [4], where similar statements were proved for edge percolation on a given graph, and a result from [10] that relates the random orientation to edge percolation. This cluster of questions on correlation of paths was inspired by an interesting conjecture due to Kasteleyn, named the bunkbed conjecture by Häggström [7]; see also [9] and Remark 5 in [3].

2. The complete graph K_n

For three different vertices, a, s and b, of K_n we want to know if the event $\{a \to s\}$ and the event $\{s \to b\}$ are positively or negatively correlated. As mentioned in the Introduction, it turns out to be easier to study the correlation of the complementary events, *i.e.*, $A := \{a \nrightarrow s\}$, that there does not exist a directed path from a to s, and the event $B := \{s \nrightarrow b\}$, which have the same correlation.

Think of the vertices of K_n as $[n] := \{1, \ldots, n\}$.

Theorem 2.1. *For all $n \geqslant 2$,*

$$\left(\frac{1}{2}\right)^{n-2}\left(1 - \left(\frac{1}{2}\right)^{n-1}\right) \leqslant \mathbb{P}(A) \leqslant \left(\frac{1}{2}\right)^{n-2}\left(1 + 3.2 \cdot \left(\frac{7}{8}\right)^{n-1}\right).$$

In particular,

$$\lim_{n\to\infty} 2^{n-2} \cdot \mathbb{P}(A) = 1.$$

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

Proof. For a vertex $x \in [n]$, let O_x denote the sets of vertices in $[n] \setminus \{x\}$ that can be reached from x in one step. Similarly, let I_x denote the sets of vertices in $[n] \setminus \{x\}$ that can reach x in one step. We want to prove that $O_a = \emptyset$ and $I_s = \emptyset$ are the two main obstructions for a path from a to s. The lower bound is easy:

$$\mathbb{P}(A) \geqslant \mathbb{P}((O_a = \emptyset) \cup (I_s = \emptyset)) = \left(\frac{1}{2}\right)^{n-1} + \left(\frac{1}{2}\right)^{n-1} - \left(\frac{1}{2}\right)^{2n-3}.$$

For the upper bound, we estimate $\mathbb{P}(A)$ with the probability that there is no path from a to s of length 1, 2 or 3. This is equivalent to $a \in O_s$, $O_a \subset O_s \setminus \{a\}$ and the event that the vertices in O_a have no directed edges to vertices in I_s, respectively. Note that, if $k = |O_a|$ and $m = |I_s|$, then $k + m \leqslant n - 2$ is necessary, and there are $k \cdot m$ edges between O_a and I_s. We separate the two terms $k = 0$ and $m = 0$ and show that these are dominant. Thus

$$\begin{aligned}
\mathbb{P}(A) &\leqslant \frac{1}{2} \cdot \sum_{k=0}^{n-2} \binom{n-2}{k} \left(\frac{1}{2}\right)^{n-2} \sum_{m=0}^{n-2-k} \binom{n-2-k}{m} \left(\frac{1}{2}\right)^{n-2} \cdot \left(\frac{1}{2}\right)^{k \cdot m} \\
&= \left(\frac{1}{2}\right)^{2n-3} \sum_{m=0}^{n-2} \binom{n-2}{m} + \left(\frac{1}{2}\right)^{2n-3} \sum_{k=1}^{n-2} \binom{n-2}{k} \\
&\quad + \left(\frac{1}{2}\right)^{2n-3} \sum_{k=1}^{n-2} \binom{n-2}{k} \sum_{m=1}^{n-2-k} \binom{n-2-k}{m} \left(\frac{1}{2}\right)^{k \cdot m} \\
&\leqslant \left(\frac{1}{2}\right)^{n-2} + \left(\frac{1}{2}\right)^{2n-3} \sum_{k=1}^{n-3} \binom{n-2}{k} \sum_{m=1}^{n-2-k} \binom{n-2-k}{m} \left(\frac{1}{2}\right)^{k \cdot m}.
\end{aligned}$$

By Lemma 2.2 below,

$$a(n) := \sum_{k=1}^{n-1} \binom{n}{k} \sum_{m=1}^{n-k} \binom{n-k}{m} \left(\frac{1}{2}\right)^{km} \leqslant 5.6 \cdot \left(\frac{7}{4}\right)^n,$$

so

$$\mathbb{P}(A) \leqslant \left(\frac{1}{2}\right)^{n-2} + 5.6 \cdot \left(\frac{1}{2}\right)^{2n-3} \left(\frac{7}{4}\right)^{n-2} \leqslant \left(\frac{1}{2}\right)^{n-2} \left(1 + 3.2 \cdot \left(\frac{7}{8}\right)^{n-1}\right). \qquad \square$$

Lemma 2.2. *For all $n \geqslant 0$,*

$$a(n) := \sum_{k=1}^{n-1} \binom{n}{k} \sum_{m=1}^{n-k} \binom{n-k}{m} \left(\frac{1}{2}\right)^{km} \leqslant 5.6 \cdot \left(\frac{7}{4}\right)^n.$$

Proof. As $a(0) = a(1) = 0$, we will assume that $n \geqslant 2$. We will use that

$$\left(\frac{1}{2}\right)^{km} = \left(\frac{1}{4}\right)^m \cdot \left(\frac{1}{2}\right)^{(k-2)m} \leqslant \left(\frac{1}{4}\right)^m \cdot \left(\frac{1}{2}\right)^{k-2} = 4 \cdot \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{4}\right)^m,$$

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

if $m \geqslant 1$ and $k \geqslant 2$, and split the sum into two parts, $k = 1$ and $k \geqslant 2$. Thus

$$\begin{aligned} a(n) &= n \cdot \sum_{m=1}^{n-1} \binom{n-1}{m} \left(\frac{1}{2}\right)^m + \sum_{k=2}^{n-1} \binom{n}{k} \sum_{m=1}^{n-k} \binom{n-k}{m} \left(\frac{1}{2}\right)^{km} \\ &\leqslant n \cdot \left(\frac{3}{2}\right)^{n-1} + 4 \cdot \sum_{k=2}^{n-1} \binom{n}{k} \left(\frac{1}{2}\right)^k \sum_{m=1}^{n-k} \binom{n-k}{m} \left(\frac{1}{4}\right)^m \\ &\leqslant n \cdot \left(\frac{3}{2}\right)^{n-1} + 4 \cdot \sum_{k=2}^{n-1} \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{5}{4}\right)^{n-k} \leqslant n \cdot \left(\frac{3}{2}\right)^{n-1} + 4 \cdot \left(\frac{7}{4}\right)^n. \end{aligned}$$

The lemma follows since $n \cdot \left(\frac{3}{2}\right)^{n-1} \leqslant 1.6 \cdot \left(\frac{7}{4}\right)^n$ holds for all $n \geqslant 2$. □

Remark. As $k = 1$ contributes $n \cdot \left(\frac{3}{2}\right)^{n-1}$, this gives a lower bound for $a(n)$. This is in fact the dominating term, so that $a(n)$ asymptotically is of order $p(n) \cdot \left(\frac{3}{2}\right)^n$, where $p(n)$ is a polynomial in n. By splitting the sum into three parts, we can, for example, show that $a(n) \leqslant 13.6 \cdot \left(\frac{13}{8}\right)^n$ for all $n \geqslant 0$.

To estimate $\mathbb{P}(A \cap B)$, the following lemma, in combination with Lemma 2.2, is useful.

Lemma 2.3. *For all $n \geqslant 0$,*

$$b(n) := \sum_{k=1}^{n-1} \binom{n}{k} \sum_{i=1}^{n-1-k} \binom{n-k}{i} \left(\frac{1}{2}\right)^{ki} \sum_{m=1}^{k} \binom{k}{m} \left(\frac{1}{2}\right)^{m(n-k-i)} \leqslant 4 \cdot \left(\frac{7}{4}\right)^n.$$

Proof. Note first that $b(0) = b(1) = b(2) = 0$, so that we may assume that $n \geqslant 3$. Note also that, for $k, i \geqslant 1$, $ki = (k-1)(i-1) + k + i - 1 \geqslant k + i - 1$, so that $\left(\frac{1}{2}\right)^{ki} \leqslant 2\left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^i$. This gives

$$\begin{aligned} b(n) &\leqslant 4 \cdot \sum_{k=1}^{n-1} \binom{n}{k} \cdot \left(\frac{1}{2}\right)^k \sum_{i=1}^{n-1-k} \binom{n-k}{i} \cdot \left(\frac{1}{2}\right)^i \cdot \left(\frac{1}{2}\right)^{n-k-i} \sum_{m=1}^{k} \binom{k}{m} \cdot \left(\frac{1}{2}\right)^m \\ &\leqslant 4 \cdot \sum_{k=1}^{n-1} \binom{n}{k} \cdot \left(\frac{1}{2}\right)^k \cdot 1 \cdot \left(\frac{3}{2}\right)^k = 4 \cdot \sum_{k=1}^{n-1} \binom{n}{k} \cdot \left(\frac{3}{4}\right)^k \leqslant 4 \cdot \left(\frac{7}{4}\right)^n. \end{aligned}$$

□

Theorem 2.4. *For all $n \geqslant 3$,*

$$\left(\frac{1}{2}\right)^{2n-3} \left(3 - 2\left(\frac{1}{2}\right)^{n-3}\right) \leqslant \mathbb{P}(A \cap B) \leqslant \left(\frac{1}{2}\right)^{2n-3} \left(3 + 20.8 \cdot \left(\frac{7}{8}\right)^{n-3}\right).$$

In particular,

$$\lim_{n \to \infty} 2^{2n-3} \cdot \mathbb{P}(A \cap B) = 3.$$

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

Proof. A necessary condition for $A \cap B$ is that the edge between a and s is directed from s to a, that the edge between s and b is directed from b to s and that the edge between a and b is directed from b to a. Let E denote this event, noting that $\mathbb{P}(E) = 1/8$.

Let O_a, O_s and O_b denote the sets of vertices in $[n] \setminus \{a, s, b\}$ that can be reached from a, s and b, respectively, in one step. Similarly, let I_a, I_s and I_b denote the sets of vertices in $[n] \setminus \{a, s, b\}$ that can reach a, s and b, respectively, in one step. For the lower bound, note that $E \cap (O_a = O_s = \emptyset) \Rightarrow A \cap B$, $E \cap (O_a = I_b = \emptyset) \Rightarrow A \cap B$ and $E \cap (I_s = I_b = \emptyset) \Rightarrow A \cap B$, so that

$$\begin{aligned}\mathbb{P}(A \cap B) &\geqslant \mathbb{P}((O_a = O_s = \emptyset) \cup (O_a = I_b = \emptyset) \cup (I_s = I_b = \emptyset))/8 \\ &= \left(\frac{1}{2}\right)^3 \left(3\left(\frac{1}{2}\right)^{2n-6} - 2\left(\frac{1}{2}\right)^{3n-9}\right) = \left(\frac{1}{2}\right)^{2n-3} \left(3 - 2\left(\frac{1}{2}\right)^{n-3}\right),\end{aligned}$$

as $(O_a = O_s = \emptyset) \cap (I_s = I_b = \emptyset) = \emptyset$.

For the upper bound, we note that $A \cap B \Rightarrow E \cap (O_a \subset O_s \subset O_b) \cap F$, where F denotes the event that the vertices in O_a have no directed edges to vertices in I_s and the vertices in O_s have no directed edges to vertices in I_b, so that $\mathbb{P}(A \cap B) \leqslant \mathbb{P}((O_a \subset O_s \subset O_b) \cap F)/8$. Let $k = |O_s|$, $i = |I_b|$ and $m = |O_a|$. Then $0 \leqslant k \leqslant n-3$, $0 \leqslant i \leqslant n-3-k$ and $0 \leqslant m \leqslant k$, and the direction of all $3(n-3)+3 = 3n-6$ edges connected to a, s and b are determined. Further, the event F determines the direction of $ki + m(n-3-k-i)$ edges, so that

$$\begin{aligned}\mathbb{P}(A \cap B) &\leqslant \left(\frac{1}{2}\right)^{3n-6} \sum_{k=0}^{n-3} \binom{n-3}{k} \sum_{i=0}^{n-3-k} \binom{n-3-k}{i} \left(\frac{1}{2}\right)^{ki} \sum_{m=0}^{k} \binom{k}{m} \left(\frac{1}{2}\right)^{m(n-3-k-i)} \\ &= \left(\frac{1}{2}\right)^{3n-6} \cdot (S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7),\end{aligned}$$

where the triple sum is split into seven parts:

(1) $k = 0 \Rightarrow m = 0$, (2) $k = n-3 \Rightarrow i = 0$, (3) $i = m = 0, 1 \leqslant k \leqslant n-4$,
(4) $i = 0, m \geqslant 1, 1 \leqslant k \leqslant n-4$, (5) $m = 0, i \geqslant 1, 1 \leqslant k \leqslant n-4$,
(6) $i = n-3-k, m \geqslant 1, 1 \leqslant k \leqslant n-4$,
(7) $1 \leqslant k \leqslant n-4, 1 \leqslant i \leqslant n-4-k, 1 \leqslant m \leqslant k$.

The first three cases correspond to the three cases of the lower bound,

$$\begin{aligned}S_1 = S_2 &= \sum_{i=0}^{n-3} \binom{n-3}{i} = 2^{n-3}, \\ S_3 &= \sum_{k=1}^{n-4} \binom{n-3}{k} = 2^{n-3} - 2 \leqslant 2^{n-3}.\end{aligned}$$

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

The next three can be expressed by the function a of Lemma 2.2,

$$S_4 = \sum_{k=1}^{n-4} \binom{n-3}{k} \sum_{m=1}^{k} \binom{k}{m} \cdot \left(\frac{1}{2}\right)^{m(n-3-k)} = \sum_{j=1}^{n-4} \binom{n-3}{j} \sum_{m=1}^{n-3-j} \binom{n-3-j}{m} \cdot \left(\frac{1}{2}\right)^{mj}$$
$$= a(n-3) \leqslant 5.6 \cdot \left(\frac{7}{4}\right)^{n-3},$$

$$S_5 = \sum_{k=1}^{n-4} \binom{n-3}{k} \sum_{i=0}^{n-3-k} \binom{n-3-k}{i} \cdot \left(\frac{1}{2}\right)^{ki} = a(n-3) \leqslant 5.6 \cdot \left(\frac{7}{4}\right)^{n-3},$$

$$S_6 = \sum_{k=1}^{n-4} \binom{n-3}{k} \cdot \left(\frac{1}{2}\right)^{k(n-3-k)} \sum_{m=1}^{k} \binom{k}{m} \leqslant \sum_{k=1}^{n-4} \binom{n-3}{k} \sum_{m=1}^{k} \binom{k}{m} \cdot \left(\frac{1}{2}\right)^{m(n-3-k)}$$
$$= \sum_{i=1}^{n-4} \binom{n-3}{i} \sum_{m=1}^{n-3-i} \binom{n-3-i}{m} \cdot \left(\frac{1}{2}\right)^{im} = a(n-3) \leqslant 5.6 \cdot \left(\frac{7}{4}\right)^{n-3},$$

and the last by the function b of Lemma 2.3,

$$S_7 = \sum_{k=1}^{n-4} \binom{n-3}{k} \sum_{i=1}^{n-4-k} \binom{n-3-k}{i} \cdot \left(\frac{1}{2}\right)^{ki} \sum_{m=1}^{k} \binom{k}{m} \cdot \left(\frac{1}{2}\right)^{m(n-3-k-i)}$$
$$= b(n-3) \leqslant 4 \cdot \left(\frac{7}{4}\right)^{n-3}.$$

Collecting the estimates gives the theorem. □

Theorem 2.5.

$$\lim_{n\to\infty} \frac{\mathbb{P}(A \cap B) - \mathbb{P}(A) \cdot \mathbb{P}(B)}{\mathbb{P}(A \cap B)} = \frac{1}{3}.$$

Proof. The proof follows immediately from Theorems 2.1 and 2.4 as $\mathbb{P}(B) = \mathbb{P}(A)$. □

Remark. Note that

$$\frac{\mathbb{P}(A \cap B) - \mathbb{P}(A) \cdot \mathbb{P}(B)}{\mathbb{P}(A \cap B)} = 1 - \frac{\mathbb{P}(A) \cdot \mathbb{P}(B)}{\mathbb{P}(A) \cdot \mathbb{P}(B \mid A)} = 1 - \frac{\mathbb{P}(B)}{\mathbb{P}(B \mid A)},$$

so that Theorem 2.5 can be formulated as

$$\lim_{n\to\infty} \frac{\mathbb{P}(B \mid A)}{\mathbb{P}(B)} = \frac{3}{2}.$$

Theorem 2.5 shows that the events $A = \{a \nrightarrow s\}$ and $B = \{s \nrightarrow b\}$ are positively correlated for sufficiently large n. From this it follows that the complementary events $\{a \to s\}$ and $\{s \to b\}$ are also positively correlated for sufficiently large n. It is in fact true for all $n \geqslant 5$, as the next theorem shows.

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

Theorem 2.6. *The events $A = \{a \nrightarrow s\}$ and $B = \{s \nrightarrow b\}$ are negatively correlated for $n = 3$, independent for $n = 4$ and positively correlated for $n \geqslant 5$.*

Proof. From Theorems 2.1 and 2.4 we get

$$\begin{aligned}
&\mathbb{P}(A \cap B) - \mathbb{P}(A) \cdot \mathbb{P}(B) = \mathbb{P}(A \cap B) - (\mathbb{P}(A))^2 \\
&\geqslant \left(\frac{1}{2}\right)^{2n-3} \cdot \left(3 - 2\left(\frac{1}{2}\right)^{n-3}\right) - \left\{\left(\frac{1}{2}\right)^{n-2} \cdot \left(1 + 3.2 \cdot \left(\frac{7}{8}\right)^{n-1}\right)\right\}^2 \\
&\geqslant \left(\frac{1}{2}\right)^{2n-4} \cdot \left(6 - 4 \cdot \left(\frac{1}{2}\right)^{n-3} - 1 - 6.4 \cdot \left(\frac{7}{8}\right)^{n-1} - 10.24 \cdot \left(\frac{49}{64}\right)^{n-1}\right) \\
&= \left(\frac{1}{2}\right)^{2n-4} \cdot (5 - c(n)),
\end{aligned}$$

where $c(n)$ is a decreasing function of n, with $c(8) < 5$, so that the theorem holds for $n \geqslant 8$. The remaining cases, $3 \leqslant n \leqslant 7$, are proved using the recursion formulas in Lemmas 3.1 and 3.2 in the next section. □

3. Exact recursions

For $n \geqslant 2$, $s \in [n]$ and $K \subset [n] \setminus \{s\}$, let $\{K \nrightarrow s\}$ denote the event $\{a \nrightarrow s \text{ for every } a \in K\}$ in K_n. With $|K| = k$, define

$$f(n,k) := \mathbb{P}_n(K \nrightarrow s),$$

where, in particular, $f(n,0) = 1$. Also set $f(1,0) = 1$ for convenience.

For $n \geqslant 3$ and $s, b \in [n], K \subset [n] \setminus \{s,b\}$, $s \neq b$ and $|K| = k$, define

$$g(n,k) := \mathbb{P}_n(K \nrightarrow s, s \nrightarrow b),$$

where, in particular, $g(n,0) = f(n,1)$. Also, let $g(2,0) := f(2,1) = 1/2$.

Lemma 3.1. *For $n \geqslant k+1 \geqslant 2$ we have*

$$f(n,k) = \sum_{i=0}^{n-k-1} \binom{n-k-1}{i} \frac{(2^k-1)^i}{2^{k(n-k)}} f(n-k,i).$$

Lemma 3.2. *For $n \geqslant k+2 \geqslant 3$ we have*

$$g(n,k) = \sum_{i=0}^{n-k-2} \binom{n-k-2}{i} \frac{(2^k-1)^i}{2^{k(n-k)}} g(n-k,i).$$

We trust the reader to be able to deduce the correctness of these recursions. Proofs can be found in the `arXiv` version of this paper [1]. Using the lemmas, we can recursively compute the desired probabilities $\mathbb{P}(A) = f(n,1)$ and $\mathbb{P}(A \cap B) = g(n,1)$ in K_n.

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

4. Other graphs

Proposition 4.1. *When G is the cyclic graph with n vertices, C_n, the covariance between the events $\{a \to s\}$ and $\{s \to b\}$ is at most $-\left(\frac{1}{2}\right)^{2n}$, with equality if and only if the vertices a and b are adjacent to s.* □

Proof. Let c, d and $n-c-d$ denote the distances (number of edges) between a and s, s and b, and b and a, respectively. We assume the three vertices to be distinct, so that $c, d, n-c-d \geqslant 1$. Then,

$$\mathbb{P}(a \to s) = \left(\frac{1}{2}\right)^{c} + \left(\frac{1}{2}\right)^{n-c} - \left(\frac{1}{2}\right)^{n},$$
$$\mathbb{P}(s \to b) = \left(\frac{1}{2}\right)^{d} + \left(\frac{1}{2}\right)^{n-d} - \left(\frac{1}{2}\right)^{n},$$
$$\mathbb{P}(\{a \to s\} \cap \{s \to b\}) = \left(\frac{1}{2}\right)^{c} \cdot \left(\frac{1}{2}\right)^{d} + \left(\frac{1}{2}\right)^{n},$$

where the last term corresponds to the case when there is a directed path $s \to a \to b \to s$. The proposition follows. □

For trees the situation is even simpler as there are no cycles, so there is a unique path between any two vertices. This implies that when G is a tree (or a forest), the events $\{a \to s\}$ and $\{s \to b\}$ are either independent or mutually exclusive.

It is not difficult to find graphs for which the sign of the correlation depends on which vertices are chosen as a, b, s. The smallest example is K_4 with one edge removed. If the vertices a, b, s are chosen so that $\{a, b\}$ is the missing edge, then the correlation is

$$\mathbb{P}(\{a \to s\} \cap \{s \to b\}) - \mathbb{P}(a \to s) \cdot \mathbb{P}(s \to b) = \frac{7}{16} - \left(\frac{21}{32}\right)^{2} = \frac{7}{1024} > 0.$$

If not, then the correlation is negative.

5. Open problems and conjectures

From the observations in Section 4, we make the following conjecture.

Conjecture 5.1. *For any connected graph $G = (V, E)$ and three distinct vertices a, s and b in V, if s has degree at most two, then the events $\{a \to s\}$ and $\{s \to b\}$ are independent or negatively correlated.*

Any connected simple graph $G = (V, E), |V| \geqslant 3$ belongs to (at least) one of the following classes.

I For any three distinct vertices $a, b, s \in V(G)$, the events $\{a \to s\}$ and $\{s \to b\}$ are non-positively correlated.

II The events $\{a \to s\}$ and $\{s \to b\}$ are independent for some distinct $a, b, s \in V(G)$ or they are negatively or positively correlated depending on the choice of $a, b, s \in V(G)$.

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press

III For any three distinct vertices $a, b, s \in V(G)$, the events $\{a \to s\}$ and $\{s \to b\}$ are non-negatively correlated.

We have seen that trees and cycles belong to Class I, $K_n, n \geqslant 5$ belongs to Class III and K_4 minus one edge belongs to Class II. Note that due to independent events there is some overlap between the classes; in particular, K_4 belongs to all three classes.

Conjecture 5.2. *For large n, most graphs will belong to Class II.*

In fact we guess that for n large enough, the graphs in Class I are joins (in some vague sense) of cycles and trees. It would be interesting if it were possible to characterize the graphs in Class I or Class III. We formulate the following more specific questions. Recall that outerplanar graphs are the graphs that do not have K_4 or $K_{2,3}$ as minors.

Problems 5.3. *We pose the following problems.*

(1) *Are all graphs in Class I (with $|V(G)| \geqslant 5$) outerplanar?*
(2) *For a given n, what is the smallest number k such that there exist k edges whose removal from K_n gives a graph not in Class III?*
(3) *Is it true that if G belongs to Class I, but not to Class II, then so does any connected subgraph obtained by removing one edge?*
(4) *Similarly, is it true that if G belongs to Class III, but not to Class II, then so does G plus any new edge?*
(5) *The results in* [2] *seem to suggest that Class III is larger than Class I. Is this true?*

Acknowledgement

We thank Svante Janson, Stanislav Volkov and Johan Wästlund for fruitful discussions. We also want to thank an anonymous referee for comments that have improved the exposition.

References

[1] Alm, S. E. and Linusson, S. A counter-intuitive correlation in a random tournament (early version of the present paper). `arXiv:0906.0240`
[2] Alm, S. E. and Linusson, S. (2009) Correlations for paths in random orientations of $G(n, p)$. Preprint. `arXiv:0906.0720`
[3] van den Berg, J. and Kahn, J. (2001) A correlation inequality for connection events in percolation. *Ann. Probab.* **29** 123–126.
[4] van den Berg, J., Häggström, O. and Kahn, J. (2006) Some conditional correlation inequalities for percolation and related processes. *Random Struct. Alg.* **29** 417–435.
[5] Grimmett, G. R. (1999) *Percolation*, Springer, Berlin.
[6] Grimmett, G. R. (2001) Infinite paths in randomly oriented lattices. *Random Struct. Alg.* **18** 257–266.
[7] Häggström, O. (2003) Probability on Bunkbed Graphs. In *Proc. FPSAC'03: Formal Power Series and Algebraic Combinatorics* (Linköping, Sweden). Available at `http://www.fpsac.org/FPSAC03/ARTICLES/42.pdf`.
[8] Linusson, S. (2009) A note on correlations in randomly oriented graphs. Preprint. `arXiv:0905.2881`.
[9] Linusson, S. (2010) On percolation and the bunkbed conjecture. *Combin. Probab. Comput.* **XX**.
[10] McDiarmid, C. (1981) General percolation and random graphs. *Adv. Appl. Probab.* **13** 40–60.

https://doi.org/10.1017/S0963548310000131 Published online by Cambridge University Press