

Random Threshold Digraphs

Elizabeth Reilly

Johns Hopkins University Applied Physics Laboratory
Laurel, Maryland 20723 USA
Elizabeth.Reilly@jhuapl.edu

Edward Scheinerman Yiguang Zhang

Applied Mathematics & Statistics Department
Johns Hopkins University
Baltimore, MD 21218 USA
ers@jhu.edu yzhan132@jhu.edu

Submitted: Jan 22, 2014; Accepted: May 26, 2014; Published: Jun 9, 2014
Mathematics Subject Classification 05C62, 05C80

Abstract

This paper introduces a notion of a *random threshold directed graph*, extending the work of Reilly and Scheinerman in the undirected case and closely related to random Ferrers digraphs.

We begin by presenting the main definition: D is a *threshold digraph* provided we can find a pair of weighting functions $f, g : V(D) \rightarrow \mathbb{R}$ such that for distinct $v, w \in V(D)$ we have $v \rightarrow w$ iff $f(v) + g(w) \geq 1$. We also give an equivalent formulation based on an *order representation* that is purely combinatorial (no arithmetic). We show that our formulations are equivalent to the definition in the work of Cloteaux, LaMar, Moseman, and Shook in which the focus is on the degree sequence, and present a new characterization theorem for threshold digraphs.

We then develop the notion of a *random threshold digraph* formed by choosing vertex weights independently and uniformly at random from $[0, 1]$. We show that this notion of a random threshold digraph is equivalent to a purely combinatorial approach, and present a formula for the probability of a digraph based on counting linear extensions of an auxiliary partially ordered set.

We capitalize on this equivalence to develop exact and asymptotic properties of random threshold digraphs such as the number of vertices with in-degree (or out-degree) equal to zero, domination number, connectivity and strong connectivity, clique and independence number, and chromatic number.

Keywords: threshold digraph, random graph

1 Introduction

1.1 Notation

In this paper *graph* exclusively means a simple, undirected graph without loops or multiple edges. We write $u \sim v$ to denote that uv is an edge.

The term *digraph* means a strict directed graph without loops or multiple edges in the same direction, but antiparallel edges ($u \rightarrow v$ and $v \rightarrow u$) may be present. Occasionally we permit loops in directed graphs in which case we use the phrase *looped digraphs*—in that case there can be at most one loop incident with a vertex.

We write $|G|$ to denote the number of vertices and $\|G\|$ the number of edges in the [di]graph G .

For a digraph D , its *underlying simple graph* is denoted $\text{simp}(D)$. That is, $\text{simp}(D)$ is a simple graph with the same vertex set as D in which $u \sim v$ if and only if $u \rightarrow v$ or $v \rightarrow u$ in D .

If \mathcal{X} is some property of a random graph [or digraph] on n vertices, we say that \mathcal{X} holds *with high probability* provided the probability \mathcal{X} holds tends to 1 as $n \rightarrow \infty$.

For integers $n \geq k \geq 0$ we use this notation for rising and falling factorial:

$$n^{\bar{k}} = n(n+1)(n+2)\cdots(n+k-1) \quad \text{and} \quad n^{\underline{k}} = n(n-1)(n-2)\cdots(n-k+1).$$

1.2 Background: Undirected (random) threshold graphs

The concept of a threshold graph originates with the following problem [4]. For a subset of vertices of a graph, $A \subseteq V(G)$, let $\mathbf{1}_A$ denote the characteristic vector of that set. That is, $\mathbf{1}_A \in \mathbb{R}^n$ (where $n = |G|$) is a 0-1 vector with 1s corresponding to elements of A and 0s in the other coordinates. The question is, can one separate subsets of vertices that are independent (induce no edges) from those that are not by a linear inequality? That is, we seek a vector \mathbf{w} and a threshold t such that

$$A \text{ is an independent set of vertices} \iff \mathbf{1}_A \cdot \mathbf{w} < t.$$

A graph for which such a linear inequality exists is called a *threshold graph*.

However, for our purposes, we use a different, but equivalent, definition. See [8] and [10].

Definition 1 (Threshold graph). A graph G is called a *threshold graph* provided there is a function $f : V(G) \rightarrow \mathbb{R}$ and a real number t such that for distinct vertices v, w we have

$$v \sim w \iff f(v) + f(w) \geq t. \tag{1}$$

There is no loss of generality in assuming that the values of the weighting function f be restricted to the unit interval $[0, 1]$ and taking the threshold t to be 1. Furthermore, by gently increasing the values assigned to all vertices, we may assume that equality never holds in (1).

The values assigned to vertices are often referred to as *weights* and the function f is called a *threshold representation* of G .

There are a variety of interesting results about threshold graphs but we only mention the following characterizations; full details in [10].

Theorem 2. *Let G be a simple graph. The following are equivalent.*

1. G is a threshold graph.
2. G does not contain any of $2K_2$, C_4 , or P_4 as an induced subgraph.
3. Assuming $n = |G| \geq 2$, G contains a vertex v with $d(v) = 0$ or $d(v) = n - 1$ and $G - v$ is a threshold graph.
4. G is uniquely determined by its degree sequence. That is, if H is a graph with $V(H) = V(G)$ and $\forall v \in V(G)$, $d_G(v) = d_H(v)$, then $H = G$. \square

An alternative perspective is to begin with a list of weights, and from this list of weights produce a threshold graph. To this end, let \mathcal{G}_n denote the set of all graphs G with vertex set $V(G) = [n] = \{1, 2, \dots, n\}$. For a vector $\mathbf{x} \in [0, 1]^n$ let $T(\mathbf{x})$ denote the graph $G \in \mathcal{G}_n$ in which, for distinct vertices u and v , we have $u \sim v$ iff $x_u + x_v \geq 1$. Thus T is a mapping from the cube $[0, 1]^n$ to the set of all n -vertex labelled graphs \mathcal{G}_n . Of course, $T(\mathbf{x})$ is, necessarily, a threshold graph.

We use the transformation T to create a natural model of random threshold graphs. One simply selects a point in $\mathbf{x} \in [0, 1]^n$ uniformly at random and apply T to yield a random threshold graph; see [11] and [12]. Specifically, we define the random graph \mathbf{G}_n to be a random variable taking values in \mathcal{G}_n such that

$$\Pr[\mathbf{G}_n = G] = \mu [T^{-1}(G)] \tag{2}$$

where μ is Lebesgue measure. Of course, if G is not a threshold graph, $\Pr[\mathbf{G}_n = G] = 0$.

In [11, 12] an equivalent, purely combinatorial model for random threshold graphs—based on creation sequences—provides a route to proving various exact and asymptotic results about random threshold graphs.

For example, we find that for a threshold graph $G \in \mathcal{G}_n$, the expression in (2) may be replaced by this formula:

$$\Pr[\mathbf{G}_n = G] = \frac{|\text{Aut}(G)|}{2^{n-1}n!} \tag{3}$$

where $\text{Aut}(G)$ is the automorphism group of G .

These combinatorial methods lead to exact results on properties of random threshold graphs. For example:

$$\Pr[\mathbf{G}_n \text{ is Hamiltonian}] = \frac{1}{2^{n-1}} \binom{n-2}{\lfloor (n-2)/2 \rfloor} \sim \frac{1}{\sqrt{2\pi n}}.$$

1.3 Overview of results

The goal of this paper is to extend the work on random threshold graphs [12] (see also [11]) to directed graphs.

(See also [11] and [13] which consider a variant of the random threshold model for bipartite graphs; such graphs are known as *random difference graphs*. We also refer the reader to [7] for another approach to random threshold graphs in the context of graph limits.)

In Section 2 we present a definition of threshold digraph that is a natural extension of Definition 1. Our definition is quite different from that of [5], but we prove that the two definitions are equivalent. Indeed we show that these two definitions are, in turn, equivalent to a third definition that is purely combinatorial. We close with a new characterization theorem that links all three definitions.

In Section 3 we introduce a model of *random* threshold digraphs based on choosing a pair of random weights for each vertex. Working directly from this definition is difficult, so we are fortunate to be able to present an equivalent, purely combinatorial model based on random orders. Calculation of the probability of a digraph becomes equivalent to counting the number of linear extensions in an auxiliary partially ordered set.

We capitalize on this alternative formulation in Section 4 to derive various exact and asymptotic properties of random threshold digraphs such as these:

- the probability of an edge from one vertex to another, the expected value and variance of the number of edges, and the asymptotic normality of the number of edges (suitably rescaled);
- exact results on the number of vertices with in-degree [resp. out-degree] equal to zero, an asymptotic result on the number of isolated vertices, and exact result on the probability there are no vertices with in-degree or out-degree equal to zero;
- with high probability random threshold digraphs are connected, but strongly connected only with probability asymptotic to $1/4$ (in which case their diameter is at most 4);
- asymptotic results on clique and independence number; and
- removing directions from a threshold digraph does not give a threshold graph, but does give a perfect graph, and this result enables us to give tight bounds on the chromatic number.

2 Threshold digraphs

2.1 Background on threshold digraphs

The authors of [5] define threshold digraphs to be those directed graphs that are uniquely determined by their *dual degree sequence*. That is, let D be a digraph and let $v \in V(D)$.

The *out degree* of v is the number of edges emanating from v and the *in degree* is the number of edges entering v ; these are denoted $d^+(v)$ and $d^-(v)$ respectively. To say that D is uniquely determined by its dual degree sequence means that if there is another digraph D' with $V(D') = V(D)$ for which

$$\forall v \in V(D), d_D^+(v) = d_{D'}^+(v) \text{ and } d_D^-(v) = d_{D'}^-(v)$$

then $D' = D$. The authors of [5] take this uniqueness property as their definition of directed threshold graphs. This is analogous to the final conclusion of Theorem 2.

The main result of [5] includes a forbidden subgraph characterization involving the following two ideas:

- An *induced three-cycle*, \vec{C}_3 , is a directed three cycle in which none of the reverse arcs are present. That is, a subgraph consisting of three distinct vertices a, b, c in which we have the arcs $a \rightarrow b$, $b \rightarrow c$, and $c \rightarrow a$, but no other arcs involving these three vertices. See the left portion of Figure 1.

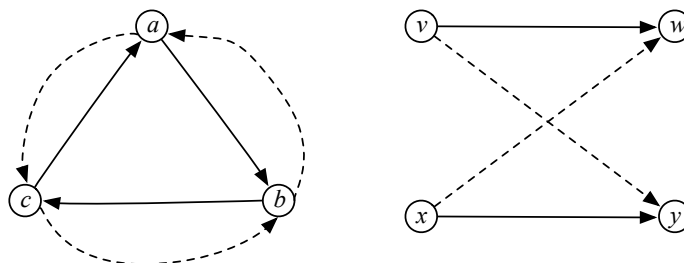


Figure 1: An induced three-cycle \vec{C}_3 (left) and a two-switch (right). In these diagrams, edges that must be present are drawn solid, edges that must be absent are drawn dashed, and all other edges are optional.

- A *two-switch* is a subgraph consisting of four distinct vertices v, w, x, y in which the edges $v \rightarrow w$ and $x \rightarrow y$ are present, the edges $v \rightarrow y$ and $x \rightarrow w$ are absent, and no conditions are placed on any other edges involving these four vertices. See the right portion of Figure 1.

It is not difficult to see that if a digraph contains either of these structures then it is not uniquely determined by its dual degree sequence. A central result in [5] is that the converse holds.

Theorem 3. *A digraph is uniquely determined by its dual degree sequence if and only if it contains neither an induced three-cycle or a two-switch.* \square

2.2 Threshold representation definition

The focus of [5] concerns matters of degree sequences and no consideration is given in that paper to representing digraphs by assigning weights to vertices. That is the approach we take here.

Rather than assigning a single weight to a vertex, we give each vertex a pair of weights: one representing a “send” weight and another being a “receive” weight. (See [16] for an analogous approach for random dot product digraphs.) There is a directed edge from one vertex to another exactly when the sum of former’s send weight and the latter’s receive weight is large enough. Here is the formal definition.

Definition 4 (Threshold digraph). Let D be a digraph. We say that D is a *threshold digraph* provided there is a pair of functions $f, g : V(D) \rightarrow [0, 1]$ such that for distinct $u, v \in V(D)$ we have

$$u \rightarrow v \iff f(u) + g(v) \geq 1. \tag{4}$$

The pair (f, g) is called a *threshold representation* of D . As in the undirected case, there is no loss of generality in assuming that equality does not hold in (4):

$$u \rightarrow v \iff f(u) + g(v) > 1 \quad \text{and} \quad u \not\rightarrow v \iff f(u) + g(v) < 1.$$

In this case, we say that the threshold representation is *strict*.

We check that if D is a threshold digraph then D does not contain an induced three-cycle or a two switch. Let (f, g) be a threshold representation of D .

- If D contains an induced three-cycle with vertex set a, b, c then we have these inequalities

$$\begin{array}{ll} f(a) + g(b) \geq 1 & f(b) + g(a) < 1 \\ f(b) + g(c) \geq 1 & f(c) + g(b) < 1 \\ f(c) + g(a) \geq 1 & f(a) + g(c) < 1 \end{array}$$

giving conflicting inequalities for $f(a) + f(b) + f(c) + g(a) + g(b) + g(c) \Rightarrow \Leftarrow$

- Similarly, if D contains a two-switch with vertex set v, w, x, y , then we have

$$\begin{array}{ll} f(v) + g(w) \geq 1 & f(v) + g(y) < 1 \\ f(x) + g(y) \geq 1 & f(x) + g(w) < 1 \end{array}$$

giving conflicting inequalities for $f(v) + f(x) + g(w) + g(y) \Rightarrow \Leftarrow$

In section 2.4 we show that the converse is also true: if D has neither an induced three-cycle nor a two-switch, then it is a threshold digraph (our definition).

2.3 Order representations and auxiliary bipartite digraphs

Our characterization of threshold digraphs and later analysis of random threshold digraphs relies on the following ideas.

Let (f, g) be a strict threshold representation of a threshold digraph D . Observe:

$$\begin{aligned} v \rightarrow w &\iff f(v) + g(w) > 1 &\iff f(v) > 1 - g(w) \\ v \not\rightarrow w &\iff f(v) + g(w) < 1 &\iff f(v) < 1 - g(w). \end{aligned}$$

Now if we make the following substitutions:

$$\hat{f}(v) = f(v) \quad \text{and} \quad \hat{g}(v) = 1 - g(v)$$

then we have $v \rightarrow w \iff \hat{f}(v) > \hat{g}(w)$. We use this as a basis for the following definition.

Definition 5. Let D be a directed graph and let L be a linearly ordered set. An *order representation* of D is a pair of functions $\hat{f}, \hat{g} : V(D) \rightarrow L$ such that for all distinct $v, w \in V(D)$ we have $v \rightarrow w \iff \hat{f}(v) > \hat{g}(w)$ and $v \not\rightarrow w \iff \hat{f}(v) < \hat{g}(w)$.

Note: We may assume $L = [0, 1]$.

This transformation gives the following.

Proposition 6. *A digraph is a threshold digraph if and only if it has an order representation.* □

This modest alternative characterization of threshold digraphs is key to our subsequent analysis of random threshold digraphs.

Our next step in showing the equivalence of our concept of threshold digraph with that of [5] is the following definition.

Definition 7. Let D be a digraph with vertex set V . Let V' and V'' be disjoint copies of V . That is, to each $v \in V$, there is a $v' \in V'$ and a $v'' \in V''$ (and $V' \cap V'' = \emptyset$).

The *auxiliary bipartite digraph* of D is a digraph B_D with vertex set $V' \cup V''$ in which we have exactly these directed edges:

- If $v \rightarrow w$ in D , then $v' \rightarrow w''$ in B_D .
- If $v \not\rightarrow w$ in D , then $v' \leftarrow w''$ in B_D .

There is never an edge, in either direction, between v' and v'' for any v . See Figure 2.

Note that if D has n vertices, then B_D has $n(n - 1)$ edges as there is always exactly one edge between v' and w'' for all $v \neq w$; it is only the direction of the edge between v' and w'' that is determined by the presence (or absence) of $v \rightarrow w$ in D .

Observe that for a threshold digraph D , the auxiliary bipartite digraph B_D encapsulates the order representation:

$$\begin{aligned} v \rightarrow w &\iff \hat{f}(v) > \hat{g}(w) &\iff v' \rightarrow w'' \\ v \not\rightarrow w &\iff \hat{f}(v) < \hat{g}(w) &\iff v' \leftarrow w'' \end{aligned}$$

Of course, B_D is defined for arbitrary digraphs D , so we need to sort out the difference.

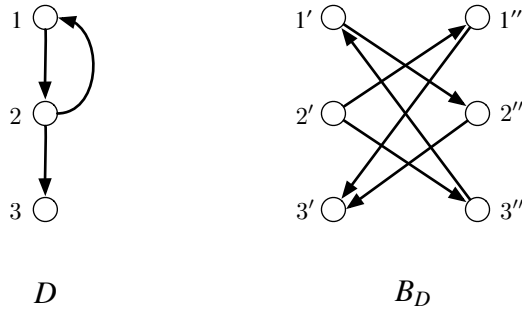


Figure 2: A digraph D and its auxiliary bipartite digraph B_D .

2.4 Characterization theorem

We now draw together the various notions we have been considering: threshold representation, order representation, auxiliary bipartite digraph, and forbidden structures.

Theorem 8. *Let D be a digraph. The following are equivalent:*

1. D is a threshold digraph.
2. D has an order representation.
3. The auxiliary bipartite digraph B_D has no directed cycles.
4. D does not contain an induced three-cycle or two-switch.

Proof. The equivalence of (1) and (2) was already noted in Proposition 6 and we have previously shown (1) \Rightarrow (4). Thus it suffices to show (4) \Rightarrow (3) \Rightarrow (2).

(4) \Rightarrow (3): Suppose D does not contain an induced three-cycle or a two-switch but, for contradiction, B_D has a shortest directed cycle

$$C = v'_1 \rightarrow v''_2 \rightarrow v'_3 \rightarrow v''_4 \rightarrow \cdots \rightarrow v'_1.$$

Since B_D can only contain at most one of $u' \rightarrow w''$ or $u' \leftarrow w''$, the cycle C has at least four vertices. Consider the following cases:

- **C has exactly four vertices:** $v'_1 \rightarrow v''_2 \rightarrow v'_3 \rightarrow v''_4 \rightarrow v'_1$.

Note that v_1, v_2, v_3, v_4 must be distinct in D . This implies that in D we have

$$v_1 \rightarrow v_2, \quad v_3 \not\rightarrow v_2, \quad v_3 \rightarrow v_4, \quad \text{and} \quad v_1 \not\rightarrow v_4$$

and therefore v_1, v_2, v_3, v_4 give a two-switch in D . $\Rightarrow \Leftarrow$

- **C has exactly six vertices:** $v'_1 \rightarrow v''_2 \rightarrow v'_3 \rightarrow v''_4 \rightarrow v'_5 \rightarrow v''_6 \rightarrow v'_1$.

– Subcase: *These six B_D vertices correspond to exactly three vertices in D .*

In this case, the cycle C must be of the form

$$a' \rightarrow b'' \rightarrow c' \rightarrow a'' \rightarrow b' \rightarrow c'' \rightarrow a'$$

where a, b, c are distinct. This implies that in D we have $a \rightarrow b \rightarrow c \rightarrow a$ but $a \not\rightarrow c \not\rightarrow b \not\rightarrow a$; that is, D contains an induced three-cycle. $\Rightarrow \Leftarrow$

– Subcase: *These six B_D vertices correspond to four or more vertices of D .*

In this case, we have $v_1 \neq v_4$, $v_2 \neq v_5$, or $v_3 \neq v_6$; without loss of generality, say $v_1 \neq v_4$. Therefore either $v'_1 \rightarrow v''_4$ or $v'_1 \leftarrow v''_4$ in B_D . In the first case B_D has the cycle $v'_1 \rightarrow v''_4 \rightarrow v'_5 \rightarrow v''_6 \rightarrow v'_1$ and in the second case B_D has the cycle $v'_1 \rightarrow v''_2 \rightarrow v'_3 \rightarrow v''_4 \rightarrow v'_1$, contradicting the fact that C is a shortest directed cycle in B_D . $\Rightarrow \Leftarrow$

• **C has exactly eight or more vertices.**

In this case $v_1 \neq v_4$ or $v_1 \neq v_6$. This implies that one of $v'_1 \rightarrow v''_4$, $v'_1 \leftarrow v''_4$, $v'_1 \rightarrow v''_6$, or $v'_1 \leftarrow v''_6$ is an edge of B_D and, as before, we find a directed cycle shorter than C . $\Rightarrow \Leftarrow$

Therefore B_D does not contain a directed cycle.

(3) \Rightarrow (2): Suppose B_D has no directed cycle.

Since B_D is a directed acyclic graph, let L be a topological sort of its vertices. That is, L is a linear order with the property that if $\alpha \rightarrow \beta$ in B_D , then $\alpha > \beta$ in L .

Define functions $\hat{f}, \hat{g} : V(D) \rightarrow L$ by

$$\hat{f}(v) = v' \quad \text{and} \quad \hat{g}(v) = v''$$

and observe that this is an order representation of D . □

3 Random model

We now develop a model of random threshold digraphs based on random representations, but also provide an equivalent, purely combinatorial formulation.

In general, a random threshold digraph is a random variable \mathbf{D}_n taking values in \mathcal{D}_n : the set of all digraphs D with $V(D) = [n] = \{1, 2, \dots, n\}$.

3.1 Random representation

Let n be a positive integer. Informally, we create a random threshold digraph \mathbf{D}_n by selecting $2n$ random numbers $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_n$ chosen independently and uniformly from $[0, 1]$. We use these to construct a digraph D with $V(D) = [n]$ in which $u \rightarrow v$ exactly when $X_u + Y_v \geq 1$.

More formally, we define a mapping from the cube $[0, 1]^{2n}$ to \mathcal{D}_n and create a random threshold digraph by choosing a point from the cube. We start with this definition.

Definition 9. Let n be a positive integer. Define $T : [0, 1]^n \times [0, 1]^n \rightarrow \mathcal{D}_n$ as follows. Let $T(\mathbf{x}, \mathbf{y})$ be the digraph D with $V(D) = [n]$ and for distinct $u, v \in [n]$ we have $u \rightarrow v$ iff $x_u + y_v \geq 1$.

We have presented this representation function T as depending on two arguments $\mathbf{x}, \mathbf{y} \in [0, 1]^n$ but, of course, we may also think of it as a function of a single $2n$ -long vector formed by concatenating \mathbf{x} and \mathbf{y} , or as n points (x_i, y_i) in the unit square $[0, 1]^2$.

Next we define the random threshold digraph \mathbf{D}_n to be the random variable taking values in \mathcal{D}_n such that

$$\Pr[\mathbf{D}_n = D] = \mu [T^{-1}(D)]$$

where μ is Lebesgue measure.

For example, let D be the digraph in Figure 3. To calculate its probability, we need



Figure 3: A threshold digraph D for which the probability $\mathbf{D}_3 = D$ is $1/180$.

to find the volume of the subset of $[0, 1]^6$ consisting of those points $(x_1, x_2, x_3, y_1, y_2, y_3)$ for which all of the following hold:

$$\begin{array}{lll} x_1 + y_2 \geq 1 & x_1 + y_3 < 1 & x_2 + y_3 \geq 1 \\ x_2 + y_1 \geq 1 & x_3 + y_1 < 1 & x_3 + y_2 < 1. \end{array}$$

The infinitely patient reader can cast this problem as the sum of these three six-fold integrals:

$$\begin{aligned} \int_0^1 dy_2 \int_0^{y_2} dy_3 \int_0^{y_3} dy_1 \int_{1-y_2}^{1-y_3} dx_1 \int_{1-y_1}^1 dx_2 \int_0^{1-y_2} dx_3 &= \frac{1}{720} \\ \int_0^1 dy_2 \int_0^{y_2} dy_3 \int_0^{y_3} dy_1 \int_{1-y_2}^{1-y_3} dx_1 \int_{1-y_1}^1 dx_2 \int_0^{1-y_2} dx_3 &= \frac{1}{720} \\ \int_0^1 dy_2 \int_0^{y_2} dy_1 \int_0^{y_1} dy_3 \int_{1-y_2}^{1-y_3} dx_1 \int_{1-y_3}^1 dx_2 \int_0^{1-y_2} dx_3 &= \frac{1}{360} \end{aligned}$$

and conclude $\Pr[\mathbf{D}_3 = D] = 1/180$. There's got to be a better way!

3.2 A combinatorial approach

Defining random threshold digraphs by choosing a random threshold representation is natural, but computationally challenging. Here we show that by choosing a random *order*

representation gives an equivalent notion of random threshold digraph that is far more tractable.

The key idea is to replace the random point $(\mathbf{x}, \mathbf{y}) \in [0, 1]^{2n}$ with the random point $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$, also in $[0, 1]^{2n}$, where

$$\hat{x}_j = x_j \quad \text{and} \quad \hat{y}_j = 1 - y_j.$$

Note that if the point (\mathbf{x}, \mathbf{y}) is chosen uniformly at random in $[0, 1]^{2n}$, the same is true for $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. Furthermore, with probability one, the $2n$ values $\hat{x}_1, \dots, \hat{y}_n$ are distinct. And, of course,

$$x_u + y_v > 1 \iff \hat{x}_u > \hat{y}_v \quad \text{and} \quad x_u + y_v < 1 \iff \hat{x}_u < \hat{y}_v.$$

The cube $[0, 1]^{2n}$ can be dissected into $(2n)!$ congruent regions based on the order of the $2n$ coordinates. To find the probability \mathbf{D}_n is a particular threshold digraph D , we just count the number of distinct order representations and divide by $(2n)!$.

More formally, consider the $(2n)!$ linear orders of the symbols $1', 2', \dots, n', 1'', 2'', \dots, n''$. Let L be such a linear order and define $T(L)$ to be the digraph in \mathcal{D}_n in which $v \rightarrow w$ when $v' > w''$ (and $v \not\rightarrow w$ when $v' < w''$).

We therefore can conclude that $\Pr[\mathbf{D}_n = D]$ equals the number of linear orders L with $T(L) = D$, divided by $(2n)!$.

We can say a bit more. Recall the auxiliary bipartite digraph B_D from Definition 7.

Definition 10 (Auxiliary poset). Let D be a threshold digraph with vertex set V and let B_D be its auxiliary bipartite digraph with vertex set $V' \cup V''$. The *auxiliary partially ordered set* P_D has ground set $V' \cup V''$ and for $\alpha, \beta \in V' \cup V''$ we have $\alpha > \beta$ exactly when there is a directed path from α to β in B_D .

Note that for threshold digraphs, B_D is acyclic and therefore the relation defined in Definition 10 indeed gives a poset. See Figure 4 for an example.

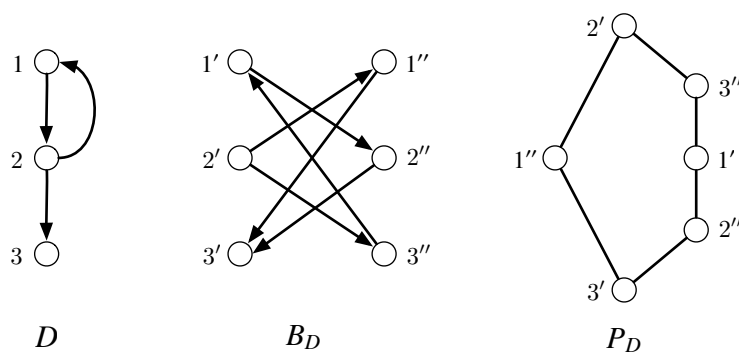


Figure 4: The auxiliary digraph B_D and auxiliary poset P_D for the digraph D from Figure 3.

We now return to the problem of computing $\Pr[\mathbf{D}_n = D]$ for a given threshold digraph D . This is tantamount to counting orders L of the symbols $1'$ through n'' such that $T(L) = D$. These orders are readily seen to be the linear extensions of P_D .

Let $\mathcal{L}(P)$ denote the set of linear extensions of a poset P . Our discussion yields the following formula.

Theorem 11. *Let n be a positive integer and let D be a threshold digraph in \mathcal{D}_n . Then*

$$\Pr[\mathbf{D}_n = D] = \frac{|\mathcal{L}(P_D)|}{(2n)!}. \quad \square$$

See Figure 5 in which this formula is used to calculate the probability \mathbf{D}_n yields the digraph from Figure 3.

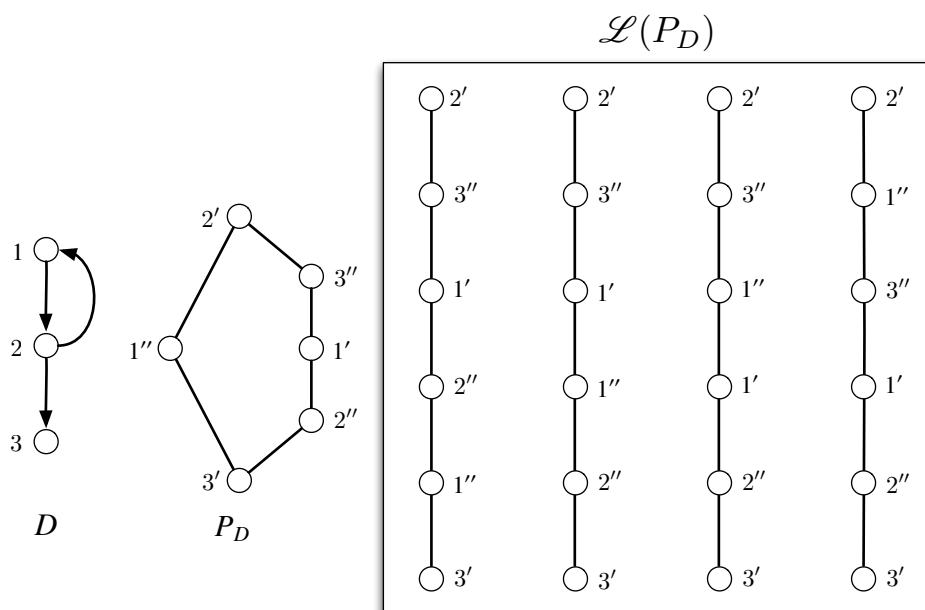


Figure 5: Computing the probability of a digraph D by counting linear extensions of P_D . Since P_D has four linear extensions $\Pr[\mathbf{D}_3 = D] = 4/6! = 1/180$.

Theorem 11 is somewhat analogous to the probability formula (3) from the undirected case. It is natural to ask if there is a way to express $\Pr[\mathbf{D}_n = D]$ as a function of just n and the size of the automorphism group of D .

The answer is no. Consider the following two threshold digraphs: the digraph D shown in Figure 3 and the directed path on three vertices: \vec{P}_3 . In both cases the digraphs have only one automorphism, but one checks that $\Pr[\mathbf{D}_n = D] = 1/180$ but $\Pr[\mathbf{D}_n = \vec{P}_3] = 1/720$.

We can extend the utility of Theorem 11 as follows.

Theorem 12. Let \mathcal{X} be any property of directed graphs. The probability \mathbf{D}_n has property \mathcal{X} equals the number of orders L of $1', 2', \dots, n', 1'', 2'', \dots, n''$ such that $T(L)$ has property \mathcal{X} , divided by $(2n)!$. \square

While the formula presented in Theorem 11 is simple and makes the problem of calculating $\Pr[\mathbf{D}_n = D]$ feasible, it is not clear how useful this formula is for large digraphs. In particular, it is known that the problem of counting the linear extensions of a general poset is #P-complete [2]. However, the posets P_D arise in a rather special way and it is conceivable that an efficient (polynomial time) algorithm for calculating $\Pr[\mathbf{D}_n = D]$ may exist. We leave this as an open question:

Question 13. What is the computational complexity of calculating the probability that \mathbf{D}_n takes on a particular value. That is, consider this problem:

- INSTANCE: A directed graph D on n vertices.
- EVALUATE: $(2n)! \Pr[\mathbf{D}_n = D]$.

Is this problem #P-complete?

Note that the corresponding problem for (undirected) threshold graphs is polynomial time computable because the calculation of $|\text{Aut}(G)|$ is easy in the case G is threshold.

3.3 Examples

Here we calculate the probability \mathbf{D}_n is an edgeless digraph and a complete bipartite digraph.

Proposition 14. Let n be a positive integer. Then the probability \mathbf{D}_n has no edges is

$$\frac{n!^2 + n!(n-1)!}{(2n)!} = \frac{1 + \frac{1}{n}}{\binom{2n}{n}} = \frac{1}{\binom{2n}{n-1}} \sim \frac{\sqrt{\pi n}}{4^n}.$$

Proof. Let $D \in \mathcal{D}_n$ be the digraph with no edges. In any order that yields D we must have $v' < w''$ for all $v \neq w$. This digraph's auxiliary poset P_D is the "standard example" poset (see [15]) shown in Figure 6. The linear extensions of this poset are of two types:

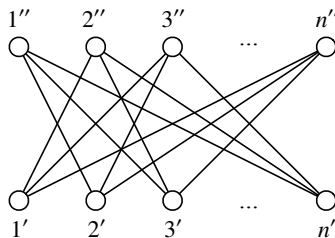


Figure 6: The poset P_D where D is an edgeless digraph with n vertices.

one in which all v' elements are below all w'' elements, and one in which a single u' is above its corresponding u'' , like this:

$$v'_1 < v'_2 < \cdots < v'_{n-1} < u'' < u' < w''_1 < w''_2 < \cdots < w''_{n-1}.$$

There are $(n!)^2$ linear extensions of the first kind and $n \cdot (n-1)!^2$ linear extensions of the second kind. The result now follows by Theorem 11. \square

For positive integers a and b , let $\overrightarrow{K_{ab}}$ be the digraph with vertex set $A \cup B$ where $A = \{1, 2, \dots, a\}$ and $B = \{a+1, a+2, \dots, a+b\}$ in which we have all edges from a vertex in A to a vertex in B , and no others.

Proposition 15. *Let a, b be positive integers and let $n = a + b$. Then*

$$\Pr \left[\mathbf{D}_n = \overrightarrow{K_{ab}} \right] = \begin{cases} \frac{[a!^2 + (a!)(a-1)!] [b!^2 + (b!)(b-1)!]}{(2a+2b)!} & \text{for } a, b > 1, \\ \frac{[(b!)^2(b+2) + (b+1)!(b-1)!]}{(2b+2)!} & \text{for } a = 1 \text{ and } b > 1, \\ 1/4 & \text{for } a = b = 1. \end{cases}$$

Proof. Figure 7 shows sample auxiliary posets for $\overrightarrow{K_{ab}}$ for the first two cases of this proposition. The result follows by counting linear extensions of these posets and applying Theorem 11. We spare the reader the gory details. \square

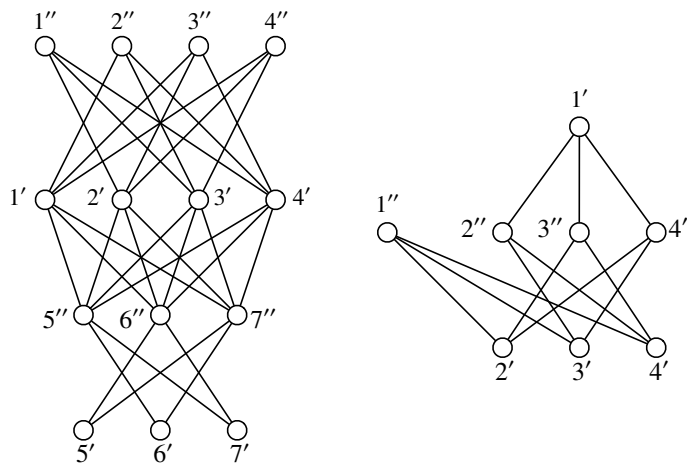


Figure 7: The auxiliary posets for the digraphs $\overrightarrow{K_{4,3}}$ and $\overrightarrow{K_{1,3}}$.

3.4 Random Ferrers digraphs

Threshold digraphs are nearly identical to Ferrers digraphs [6]; the only difference is that Ferrers digraphs have loops. See also [10].

Definition 16. Let D be a looped digraph. We say that D is a *Ferrers digraph* provided there are functions $f, g : V(D) \rightarrow [0, 1]$ such that for all $v, w \in V(D)$ we have $v \rightarrow w$ iff $f(v) + g(w) \geq 1$.

If we are given a Ferrers digraph D and remove the loops from D , the result is a threshold digraph. That raises the question: Given a threshold digraph D , how can we place loops on the vertices so that the resulting digraph is a Ferrers digraph? That is, on which vertices should we place loops to convert D into a Ferrers digraph?

One can check that the following works. Assume $V(D) = [n]$ and find all linear extensions of P_D . For a given linear extension L , place a loop on a vertex v exactly when $v' > v''$ to give a Ferrers digraph. Stepping through the various linear extensions of P_D will generate all possible Ferrers digraphs formed by adding loops to D .

For example, consider the digraph D in Figure 3; its auxiliary poset P_D has four linear extensions as shown in Figure 5. In all four extensions we have $3' < 3''$ and $2' > 2''$, but in some $1' < 1''$ and in others $1' > 1''$. This implies that there exactly two ways to add loops to D to make it a Ferrers digraph: add loops to both vertices 1 and 2, or have a loop only at vertex 2.

A study of *random* Ferrers digraphs is presented in [11]. The fundamental ideas are the same. Let \mathcal{F}_n denote the set of all looped digraphs with vertex set $[n]$. For $(\mathbf{x}, \mathbf{y}) \in [0, 1]^n \times [0, 1]^n$, let $F(\mathbf{x}, \mathbf{y})$ be the looped digraph in \mathcal{F}_n in which $i \rightarrow j$ iff $x_i + y_j \geq 1$. Define the random variable \mathbf{F}_n , taking values in \mathcal{F}_n , with $\Pr[\mathbf{F}_n = D] = \mu[F^{-1}(D)]$. In [11] a combinatorial formulation (patterned after creation sequences) provides a tool for the analysis of these random digraphs.

A simple relation connects \mathbf{F}_n and \mathbf{D}_n . For a digraph $D \in \mathcal{D}_n$, let $\ell(D)$ denote the set of all possible looped digraphs we can form by adding loops to D . Thus $\ell(D)$ contains 2^n looped digraphs. For $D \in \mathcal{D}_n$ we have

$$\Pr[\mathbf{D}_n = D] = \sum_{D' \in \ell(D)} \Pr[\mathbf{F}_n = D'].$$

(The left side implicitly counts all orders L such that $T(L) = D$; the right side counts the same orders but separates cases in which $v' < v''$ versus $v' > v''$ depending on whether or not there is a loop at v .)

Thus, if some property \mathcal{X} holds with high probability for random Ferrers digraphs, it follows that \mathcal{X} holds with high probability for random threshold digraphs.

4 Results

4.1 Edges and degrees

It is easy to check that the probability there is an edge from one vertex to another in a random threshold digraph is $\frac{1}{2}$ and that edges $u \rightarrow v$ and $v \rightarrow u$ are independent. Further, if u, v, x, y are distinct vertices, the edges $u \rightarrow v$ and $x \rightarrow y$ are independent. However, in contrast to Erdős-Rényi random graphs, there are dependencies between edges. This first result illustrates the use of Theorem 12.

Proposition 17. Let n be a positive integer and let $u, v, w \in [n]$ be distinct. In a random threshold digraph we have

$$\Pr[u \rightarrow w \mid u \rightarrow v \text{ and } v \rightarrow w] = \frac{5}{6}.$$

Proof. As mentioned, $\Pr[u \rightarrow v \wedge v \rightarrow w] = \frac{1}{4}$. To find $\Pr[u \rightarrow w \wedge u \rightarrow v \wedge v \rightarrow w]$ entails counting orders in which we have $u' > w'', u' > v'',$ and $v' > w''$. These requirements are nicely illustrated in the N-shaped poset in Figure 8. Counting the orders in which these

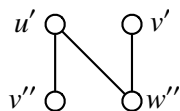


Figure 8: Relations between $u', v', v'',$ and w'' in the proof of Proposition 17.

relations hold is tantamount to counting the linear extensions of this poset: there are 5. Therefore $\Pr[u \rightarrow w \wedge u \rightarrow v \wedge v \rightarrow w] = 5/4!$ and we conclude

$$\Pr[u \rightarrow w \mid u \rightarrow v \text{ and } v \rightarrow w] = \frac{\Pr[u \rightarrow w \wedge u \rightarrow v \wedge v \rightarrow w]}{\Pr[u \rightarrow v \wedge v \rightarrow w]} = \frac{5/24}{1/4} = \frac{5}{6}. \quad \square$$

A more interesting example is afforded by calculating the variance in the number of edges in a random threshold digraph: $\text{Var} \|\mathbf{D}_n\|$.

Proposition 18. Let n be a positive integer. Then

$$\mathbb{E} \|\mathbf{D}_n\| = \frac{n^2 - n}{2} \quad \text{and} \quad \text{Var} \|\mathbf{D}_n\| = \frac{2n^3 - 3n^2 + n}{12}.$$

Proof. For distinct $i, j \in [n]$ define $X_{ij} = \mathbf{1}\{i \rightarrow j\}$, i.e., the indicator random variable that there is an edge from i to j . Let $X = \sum X_{ij}$, so $X = \|\mathbf{D}_n\|$. That $\mathbb{E} \|\mathbf{D}_n\| = \mathbb{E} X = \frac{1}{2}n(n-1)$ is immediate from $\Pr[1 \rightarrow 2] = \frac{1}{2}$. As $\text{Var} X = \mathbb{E}(X^2) - (\mathbb{E} X)^2$, we first calculate $\mathbb{E}(X^2)$.

We consider the various terms in

$$\mathbb{E}(X^2) = \sum_{i \neq j} \sum_{u \neq v} \mathbb{E}(X_{ij} X_{uv})$$

based on the overlap of the subscripts. We get the following:

Term type	Number of such terms	Expected value
$X_{12}X_{34}$	$n(n-1)(n-2)(n-3)$	$1/4$
$X_{12}X_{13}$ or $X_{21}X_{31}$	$2n(n-1)(n-2)$	$2/3!$
$X_{12}X_{31}$ or $X_{21}X_{13}$	$2n(n-1)(n-2)$	$1/4$
$X_{12}X_{21}$	$n(n-1)$	$1/4$
$X_{12}X_{12}$	$n(n-1)$	$1/2$

Note: The calculation of $E(X_{12}X_{13})$ follows from the requirement that we have $1' > 2''$ and $1' > 3''$. There are $3!$ orders for the symbols $1'$, $2''$, and $3''$ of which exactly two yield the $1 \rightarrow 2$ and $1 \rightarrow 3$.

Adding the terms in the chart and subtracting $(E X)^2$ gives the result. \square

If the edges of a random threshold digraph were mutually independent (as in Erdős-Rényi graphs) then the variance would be on the order of n^2 ; the higher variance is due to the edge dependencies. And if the edges were mutually independent, then the number of edges would converge to a normal distribution. Despite dependencies, the same holds for random threshold digraphs.

Theorem 19. *Let n be a positive integer and set $\mu_n = E \|\mathbf{D}_n\|$ and $\sigma_n^2 = \text{Var} \|\mathbf{D}_n\|$. Then*

$$\frac{\|\mathbf{D}_n\| - \mu_n}{\sigma_n}$$

converges in distribution to a standard normal random variable.

Proof. The number of edges in a random threshold digraph is readily seen to be a U-statistic and convergence to normality follows from a result of Hoeffding [9]. See Theorem 5.5.1A in [14]. \square

In an Erdős-Rényi random graph, vertex degrees are highly concentrated near the average degree. By contrast, the in- and out-degrees of vertices in \mathbf{D}_n are uniformly distributed between 0 and $n - 1$ as shown in the next proposition.

Proposition 20. *Let n, k be integers with $0 \leq k < n$. For a given vertex v , the probability $d^+(v) = k$ is $1/n$.*

Proof. Consider the out-degree of vertex 1 in \mathbf{D}_n . Consider the $n!$ orders of the n symbols $1'$ and $2'', 3'', \dots, n''$. There are $(n - 1)!$ ways to order the symbols $2'', \dots, n''$ and then exactly one way to insert $1'$ into that order so that $d^+(1) = k$. Therefore $\Pr[d^+(1) = k] = (n - 1)!/n! = 1/n$. \square

It's easy to check that, for a given vertex v , the in-degree and out-degree of v are independent. Therefore, the probability that a vertex of \mathbf{D}_n is isolated is $1/n^2$. Therefore the expected number of isolated vertices in \mathbf{D}_n is $1/n$ and Markov's inequality gives the following result.

Proposition 21. *With high probability, \mathbf{D}_n has no isolated vertices.* \square

This result suggests that, with high probability, \mathbf{D}_n is connected. We explore this issue in Section 4.3.

While it is unlikely for a vertex to have both in- and out-degree equal to zero, we may have one of these be zero and here we present some results in that regard.

We call a vertex v with out-degree $d^+(v) = 0$ a *sink* and a vertex with $d^-(v) = 0$ is called a *source*.

Proposition 22. *Let $n \geq 2$. The probability that \mathbf{D}_n has no sinks is $\frac{1}{2} - \frac{1}{2n}$. The probability it has exactly k sinks, where $1 \leq k \leq n$, is*

$$\frac{(n+1)^{k+1}}{(2n)^{k+1}}.$$

Note that if $k = n$, then \mathbf{D}_n has no edges and the formula in Proposition 22 reduces to $\binom{2n}{n-1}^{-1}$ which agrees with Proposition 14.

Proof. First we count the number of orders L of $1', \dots, n', 1'', \dots, n''$ such that the resulting digraph $T(L)$ has no sinks. This requires that for every $u \in [n]$ there is a $v \in [n]$ with $v \neq u$ such that $u' > v''$; this ensures that u has at least one out-neighbor.

Without loss of generality, we assume that elements $1''$ through n'' are in the order $1'' < 2'' < \dots < n''$; we multiply by $n!$ at the end to account for other possibilities.

Into this we can insert element $1'$ anywhere above $2''$ (otherwise $1'$ is a sink); there are $n - 1$ such choices.

Next insert elements $2'$ through n' into this list of $n + 1$ elements anywhere above $1''$ (for otherwise the inserted element would be a sink). The number of ways to add these elements to the order is $(n + 1)^{n-1}$.

Therefore, probability \mathbf{D}_n has no sink is

$$\frac{n!(n-1)(n+1)^{n-1}}{(2n)!} = \frac{1}{2} - \frac{1}{2n}.$$

Next we count orders of the $2n$ symbols $1'$ through n'' that produce a threshold digraph with exactly k sinks. As before, we only consider orders with $1'' < 2'' < \dots < n''$ and multiply by $n!$ at the end.

Let L be such an order and let S be the set of k sinks in $T(L)$. We consider the cases $1 \notin S$ and $1 \in S$ separately.

Case I: $1 \notin S$. There are $\binom{n-1}{k}$ choices for this set.

Let the $n - k$ non-sink elements be $1, u_2, u_3, \dots, u_{n-k}$. We first insert the elements $1', u'_2, u'_3, \dots, u'_{n-k}$ into the order $1'' < 2'' < \dots < n''$. Starting with $1'$, we note that we must have $1' > 2''$ for otherwise 1 would be a sink; there are $n - 1$ choices for its insertion point.

The remaining elements u'_2 through u'_{n-k} may be inserted anywhere above element $1''$ (because placing $u'_j < 1''$ would make u_j a sink); the number of ways to insert these elements is therefore $(n + 1)^{n-k-1}$.

Therefore there are $(n - 1)(n + 1)^{n-k-1}$ ways to insert the nonsinks into the order $1'' < 2'' < \dots < n''$.

Finally, we insert k symbols s' for $s \in S$. Because these are sinks, they must all be below $1''$, but in any order. This gives $k!$ choices. Therefore the number of orderings in this case ($1 \notin S$) is

$$\binom{n-1}{k} k!(n-1)(n+1)^{n-k-1}. \tag{5}$$

Case II: $1 \in S$. There are $\binom{n-1}{k-1}$ choices for this set. We separately count the orders in which $1' < 1''$ and in which $1' > 1''$.

Subcase (a): $1' < 1''$.

We have $1' < 1'' < 2'' < \dots < n''$. For $s \in S$ with $s \neq 1$ we must have $s' < 1''$ (otherwise we have the arc $s \rightarrow 1$). In other words, the elements of S appear in any order below $1''$, and there are $k!$ ways that can happen.

Having inserted all the elements s' with $s \in S$ into the order, we now insert the remaining $n - k$ elements v' . Necessarily these must lie above $1''$ (because they are not sinks) and there there are $n^{\overline{n-k}}$ ways to do this.

Thus the number of orders in this subcase is $k!n^{\overline{n-k}}$.

Subcase (b): $1' > 1''$.

Since 1 is a sink, $1'$ cannot be above $2''$ and so we must have $1'' < 1' < 2'' < 3'' < \dots < n''$. For the other elements $s \in S$ we must have $s' < 1''$ (or else s is not a sink) and these may lie there in any order. Hence there are $(k - 1)!$ ways to insert these elements.

Next we insert the elements v' for $v \notin S$ into the order. These elements must be above $1''$ (or else they are sinks) but may lie on either side of $1'$. The number of ways to insert these elements is $(n + 1)^{\overline{n-k}}$.

Thus there are $(k - 1)!(n + 1)^{\overline{n-k}}$ orders in this subcase.

We now combine these subcases, multiply by $\binom{n-1}{k-1}$ for the choices for the set S , and find that the number of orders in Case II is

$$\binom{n-1}{k-1} \left[k!n^{\overline{n-k}} + (k-1)!(n+1)^{\overline{n-k}} \right]. \quad (6)$$

Finally, we add (5) and (6), multiply by $n!$ (for the orders of $1'', 2'', \dots, n''$), divide by $(2n)!$, and simplify to get the result. \square

Corollary 23. *For all $n \geq 1$, the expected number of sinks in \mathbf{D}_n is 1. Furthermore, for a given nonnegative integer k we have*

$$\lim_{n \rightarrow \infty} \Pr [\mathbf{D}_n \text{ has exactly } k \text{ sinks}] = \frac{1}{2^{k+1}}. \quad \square$$

In section 4.3 we examine the probability that \mathbf{D}_n is strongly connected. A digraph that has a source or a sink cannot be strongly connected. Proposition 22 shows that the probability \mathbf{D}_n has no sinks is asymptotically $\frac{1}{2}$; the same is true, of course, for sources. If having no sinks and having no sources were independent, then it would be a simple matter to conclude that the probability \mathbf{D}_n has neither is asymptotically $\frac{1}{4}$. This is so, but it takes more work.

Proposition 24. *Let $n \geq 2$ be an integer. The probability that \mathbf{D}_n has no sinks and no sources is*

$$\frac{n^3 - 3n^2 + 5n - 2}{2n(2n - 1)(n - 1)} - \frac{(n - 1)!(n - 2)!}{2(2n - 1)!}$$

which converges to $\frac{1}{4}$ as $n \rightarrow \infty$.

Proof. We consider the orders for $1', \dots, n''$ that yield threshold digraphs without sources or sinks.

In such an order, the least element must not be of the form v' for then v would have no outgoing edges. Likewise, the greatest element cannot be of the form v'' . Therefore, the relevant orders must be of the form $v'' < \dots < w'$ where we might have $v = w$ or $v \neq w$; these are mutually exclusive cases that we consider separately.

In case $v = w$, the order is of the form $v'' < \dots < v'$; call this event A_1 . Notice that in this case, vertex v has edges to and from every other vertex and so there are no sources and no sinks. We have

$$\Pr(A_1) = \frac{n(2n - 2)!}{(2n)!} = \frac{1}{4n - 2}.$$

The case $v \neq w$ is more complicated. Let A_2 be the event that the order is of the form $v'' < \dots < w'$ with $v \neq w$. We have

$$\Pr(A_2) = \frac{n(n - 1)(2n - 2)!}{(2n)!} = \frac{n - 1}{4n - 2}.$$

Without loss of generality, we examine the case $1'' < \dots < n'$.

Notice for a vertex x with $1 < x < n$ we already have $x \rightarrow 1$ and $n \rightarrow x$. Thus x is neither source nor sink, 1 is not a source, and n is not a sink.

- To ensure that 1 is not a sink, we must not have

$$1' < \underbrace{2'', 3'', \dots, n''}_{\text{any order}}.$$

We call this obstruction B_1 and we have

$$\Pr(B_1) = \frac{(n - 1)!}{n!} = \frac{1}{n}.$$

- To ensure that n is not a source, we must not have

$$\underbrace{1', 2', \dots, (n - 1)'}_{\text{any order}} < n''.$$

We call this obstruction B_2 and $\Pr(B_2) = 1/n$ also.

Combining these cases we have that the probability that \mathbf{D}_n has no source and no sink is precisely $\Pr(A_1) + \Pr(A_2) \Pr(\overline{B_1 \cup B_2})$. Since $\Pr(\overline{B_1 \cup B_2}) = 1 - \Pr(B_1) - \Pr(B_2) + \Pr(B_1 \cap B_2)$ we work to calculate $\Pr(B_1 \cap B_2)$.

Since B_1 holds, we have $1' < [2'', 3'', \dots, n'']$ and we condition on the location of n'' : in this list of n symbols, it might be in position 2, 3, \dots , or n . Suppose it is in the k^{th} position:

$$1' < ?_2'' < ?_3'' < \dots < ?_{k-1}'' < n'' < ?_{k+1}'' < \dots < ?_n''.$$

The number of ways this can happen is $(n-2)!$. We now insert the symbols $2', 3', \dots, (n-1)'$ ensuring that all of these are less than n'' (as required by B_2). The number of ways to do this is $k(k+1)(k+2) \dots (k+n-3) = k^{n-2}$. Thus

$$\Pr(B_1 \cap B_2) = \frac{(n-2)!}{(2n)!} \sum_{k=2}^n k^{n-2} = \frac{(n-2)!}{(2n)!} \cdot \frac{(2n-2)! - (n-1)!^2}{(n-1)(n-1)!}.$$

Substituting these into $\Pr(A_1) + \Pr(A_2)P(\overline{B_1 \cup B_2})$ gives the results claimed. \square

4.2 Domination

Definition 25. Let D be a digraph. A set of vertices $S \subseteq V(D)$ is an *out-dominating set* provided for every vertex v either (a) $v \in S$ or (b) there is a vertex $w \in S$ such that $w \rightarrow v$.

The smallest size of an out dominating set is called the *out-domination number* of D and is denoted $\gamma^+(D)$.

Proposition 26. Let n, k be integers with $n \geq k \geq 1$. Then

$$\Pr [\gamma^+(\mathbf{D}_n) = k] = \frac{n^k}{(2n-1)^k}.$$

Proof. We count the orders of $1', \dots, n''$ that give rise to digraphs D with $\gamma^+(D) = k$. We consider the location of the last (i.e., greatest) single-primed element v' . There are two possibilities: (a) none of the double-primed elements above v' is v'' and (b) one of the double-primed elements above v' is v'' . These two cases look like this:

$$\begin{aligned} \text{(a): } & \text{other elements} < v' < \underbrace{w_1'', w_2'', \dots, w_{k-1}''}_{\text{any order}} \\ \text{(b): } & \text{other elements} < v' < \underbrace{w_1'', w_2'', \dots, w_{k-1}'', v''}_{\text{any order}} \end{aligned}$$

The number of orders of type (a) is $n^{k-1}(n-k+1)(2n-k)!$ and the number of orders of type (b) is $n^k \cdot k \cdot (2n-k-1)!$. Adding these and dividing by $(2n)!$ gives the result. \square

Corollary 27. $E[\gamma^+(\mathbf{D}_n)] = 2n/(n+1)$ and $\lim_{n \rightarrow \infty} P[\gamma^+(\mathbf{D}_n) = k] = 2^{-k}$. \square

We can consider the domination number of the underlying simple graph of random threshold digraph: $\gamma[\text{simp}(\mathbf{D}_n)]$. Computer experiments suggest that with high probability $\text{simp}(\mathbf{D}_n)$ contains a vertex of degree $n - 1$, that is, a vertex adjacent to all others. In this case, the domination number equals 1 and we pose this as an open problem.

Conjecture 28. With high probability $\text{simp}(\mathbf{D}_n)$ contains a vertex adjacent to all others.

Note that this conjecture stands in a bit of contrast to the fact that, with high probability, $\text{simp}(\mathbf{D}_n)$ does *not* contain a vertex of degree 0.

4.3 Connectivity

We have seen that the probability \mathbf{D}_n has an isolated vertex is vanishingly small (Proposition 21) and that the probability \mathbf{D}_n has no sources or sinks is asymptotically $1/4$ (Proposition 24). This suggests—and, in fact, it is the case—that with high probability \mathbf{D}_n is connected, but the probability it is strongly connected tends to $1/4$ as $n \rightarrow \infty$.

Proposition 29. *With high probability $\text{simp}(\mathbf{D}_n)$ is connected.*

Proof. Let D be a digraph with vertex set V . We know that $\text{simp}(D)$ is not connected iff there is a subset of vertices A (with $A \neq \emptyset$ and $A \neq V$) such that there are no directed edges in D between A or $V - A$ (in either direction).

For a fixed A of size a , the probability that there are no arcs from A to $V - A$ is $a!(n - a)!/n!$ because in an order representation all $v \in A$ and $w \in V - A$ we must have $v' < w''$.

Let X_A be the indicator random variable that there are no arcs between A and $V - A$. Since the arcs from A to $V - A$ are independent of the arcs from $V - A$ to A , we have $P(X_A = 1) = \binom{n}{a}^{-2}$.

Note that $\text{simp}(\mathbf{D}_n)$ is connected iff $X = \sum X_A = 0$, where the sum is over all nonempty proper subsets of V . Therefore

$$\begin{aligned} \Pr[\text{simp}(\mathbf{D}_n) \text{ is not connected}] &= \Pr(X > 0) \leq \mathbb{E}(X) = \sum_A X_A = \sum_{a=1}^{n-1} \binom{n}{a} / \binom{n}{a}^2 \\ &= \sum_{a=1}^{n-1} \frac{1}{\binom{n}{a}} = \frac{2}{n} + \sum_{a=2}^{n-2} \frac{1}{\binom{n}{a}} < \frac{2}{n} + \frac{n}{\binom{n}{2}} = O(1/n) \rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$. □

We now turn to the question of strong connectivity. There is, asymptotically, only a 25% chance that \mathbf{D}_n has no sources or sinks, and that puts an upper bound on the probability \mathbf{D}_n is strongly connected. For the lower bound, we have the following.

Proposition 30. *For $n \geq 3$, the probability that \mathbf{D}_n is strongly connected with diameter at most 4 is at least*

$$\frac{1}{4n - 2} + \frac{n - 1}{4n - 2} \left(1 - \frac{H_{n-1}}{n - 1} \right)$$

where $H_{n-1} = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n-1}$.

Proof. We consider which of the $(2n)!$ orders of $1'$ through n'' give rise to a strongly connected digraph.

Note that if the first element of the order is of the form v' , then v can have no outgoing arcs. Likewise, if the last element of the order is of the form v'' , then v can have no incoming arcs. In either case, the digraph is not strongly connected. Therefore, we need only consider orders of one of the following two forms:

$$\begin{cases} v'' < \dots < v' & \text{or} \\ v'' < \dots < w' & \text{where } v \neq w \end{cases}$$

Note that the probability that the order is of the first type is $n(2n - 2)!/(2n)! = 1/(4n - 2)$. In all such cases there are arcs from v to all other vertices and arcs from all other vertices to v , and so the digraph is strongly connected (and has diameter 2).

The probability the order is of the second form is $n(n - 1)(2n - 2)!/(2n)! = (n - 1)/(4n - 2)$. Without loss of generality, we assume we have $1'' < \dots < 2'$. In this case, for all other vertices u we have $2 \rightarrow u \rightarrow 1$. If there is at least one path of the form $1 \rightarrow v \rightarrow 2$ then the digraph is strongly connected with diameter at most 4. See Figure 9.

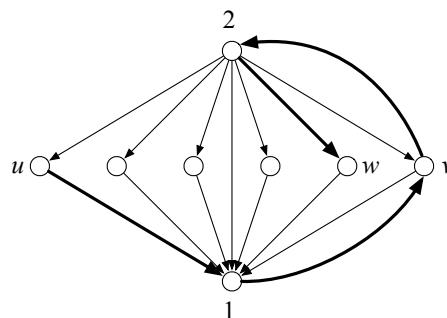


Figure 9: We are given that the threshold digraph arises from an order of the form $1'' < \dots < 2'$. This implies that $2 \rightarrow v$ and $v \rightarrow 1$ holds for all vertices $v > 2$. If there is a path of the form $1 \rightarrow v \rightarrow 2$ for $v > 2$, then the digraph must be strongly connected with diameter at most 4. One path between arbitrary vertices u and w is shown.

What is the probability there are no paths of the form $1 \rightarrow v \rightarrow 2$ with $v > 2$? We note that this depends on the order of the various symbols *other than* $1''$ and $2'$: we simply cannot have $v'' < 1'$ and $2'' < v'$ for any $v > 2$. Hence the event

$$A = \{\text{no path of the form } 1 \rightarrow v \rightarrow 2 \text{ for } v > 2\}$$

is independent of the event that $1''$ is least and $2'$ is greatest in the order.

Rather than considering the $(2n - 2)!$ orders for the elements involved in event A , we instead consider the order of $2n - 2$ iid, uniform $[0, 1]$ random values X_1, Y_2 and X_v, Y_v for $v > 2$. We then seek the probability that we do not have both $X_1 > Y_v$ and $X_v > Y_2$ for

any $v > 2$. Conditioning on $X_1 = x$ and $Y_2 = y$, the probability that $x > Y_v$ and $X_v > y$ is $x(1 - y)$. Hence

$$\Pr(A) = \int_0^1 \int_0^1 [1 - x(1 - y)]^{n-2} dx dy$$

which evaluates to $H_{n-1}/(n - 1)$ and the result follows. □

Corollary 31.

$$\lim_{n \rightarrow \infty} \Pr[\mathbf{D}_n \text{ is strongly connected}] = \frac{1}{4}. \quad \square$$

Our analysis shows that, with high probability, if \mathbf{D}_n is strongly connected, then its diameter is at most 4. Conceivably, a stronger result (lower diameter) may be achievable. However, computer experiments suggest that when \mathbf{D}_n is connected, its diameter is either 3 or 4. (Ruling out diameter 1 is easy.) These computer experiments lead us to posit the following.

Conjecture 32. For a positive integer k we have

$$\lim_{n \rightarrow \infty} \Pr[\text{diam}(\mathbf{D}_n) = k \mid \mathbf{D}_n \text{ is strongly connected}] = \begin{cases} 3/4 & \text{for } k = 3, \\ 1/4 & \text{for } k = 4, \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

In the theory of Erdős-Rényi random graphs, connectivity and Hamiltonicity are closely linked. It is reasonable to ask: What is the probability that \mathbf{D}_n has a directed Hamiltonian cycle. And a reasonable guess is: $\frac{1}{4}$ (the same as the probability of strong connectivity). However, [11] shows that the probability that a random Ferrers digraph has a directed Hamiltonian cycle tends to 0 as $n \rightarrow \infty$. Therefore, with high probability \mathbf{D}_n does not have a directed Hamiltonian cycle. See [11] for details.

4.4 Chromatic number

In this section we show that the chromatic number of \mathbf{D}_n is, with high probability, asymptotically $n/2$. The definition of the chromatic number of a digraph is the same as for simple graphs; that is, $\chi(D) = \chi[\text{simp}(D)]$.

Our method is in two steps.

First, we show that the clique number of $\text{simp}(\mathbf{D}_n)$ is, with high probability, asymptotically $n/2$. We use the same ideas to show that the independence number is asymptotically $n/4$.

Second, we show that the underlying simple graphs of threshold digraphs are perfect. It is known (and easy to prove) that threshold graphs are perfect, so one might hope that the underlying simple graph of a threshold digraph is a threshold graph, but that is not the case. Here's an example. Let D be the digraph with vertex set $[4]$ and directed edges $1 \rightarrow 2 \leftarrow 3 \rightarrow 4$. Note that D is a threshold digraph and here is a representation:

v	1	2	3	4
$f(v)$	0.4	0.4	0.7	0.1
$g(v)$	0.2	0.8	0.5	0.5

Note that even though D is a threshold digraph, $\text{simp}(D)$ is not a threshold graph.

We begin with the following result concerning the clique number of the underlying simple graph of a random threshold digraph. Note that we use $\text{CL}(\cdot)$ to denote the size of a largest clique in a graph and use $\omega(1)$ to denote a function of n that tends to infinity as $n \rightarrow \infty$.

Proposition 33. *With high probability $\text{CL}[\text{simp}(\mathbf{D}_n)] \sim n/2$. Specifically, with high probability we have*

$$\frac{n}{2} - \omega(1)\sqrt{n} \leq \text{CL}[\text{simp}(\mathbf{D}_n)] \leq \frac{n}{2} + \omega(1)\sqrt{n}.$$

Proof. Create a random order L of $1', 1'', \dots, n', n''$ and let $\mathbf{D}_n = T(L)$. Let $G = \text{simp}(\mathbf{D}_n)$.

Partition $V(\mathbf{D}_n) = V(G) = [n]$ as follows:

$$A = \{v : v' > v''\} \quad \text{and} \quad B = \{v : v' < v''\}.$$

Note that $|A|$ is a binomial $(n, \frac{1}{2})$ random variable and so

$$\frac{n}{2} - \omega(1)\sqrt{n} < |A| < \frac{n}{2} + \omega(1)\sqrt{n}$$

with high probability.

Note that $\text{CL}(G[A]) \leq \text{CL}(G) \leq \text{CL}(G[A]) + \text{CL}(G[B])$.

We claim that A is a clique. Let $v, w \in A$. Without loss of generality $v' < w'$. It follows that $w' > v' > v''$ and so $w \rightarrow v$ in \mathbf{D}_n , ergo $w \sim v$ in G . Therefore

$$\text{CL}(G) \geq \text{CL}(G[A]) = |A| > \frac{n}{2} - \omega(1)\sqrt{n}.$$

For the upper bound, it is enough to show that $\text{CL}(G[B]) = O(\sqrt{n})$ with high probability.

Let K be a k -element subset of B ; say $K = \{v_1, v_2, \dots, v_k\}$ and let X_K be the indicator random variable that K is a clique in $G[B]$. This requires that $v_i' < v_i''$ for all $i \in K$ (because $K \subseteq B$). We also claim that no v_i' (or v_i'') can lie between v_j' and v_j'' in the order. That is, neither of these is possible:

$$v_j' < v_i' < v_j'' \quad \text{or} \quad v_j' < v_i'' < v_j''. \quad (7)$$

Here's why:

In the first case, because $v_i' < v_j''$ we have $v_i \not\rightarrow v_j$. But then because $v_i'' > v_i'$ we have $v_i'' > v_j'$ which implies that $v_j \not\rightarrow v_i$. Therefore $v_i \not\sim v_j$ contradicting that K is a clique.

In the second case, because $v''_i > v'_j$ we have $v_j \not\rightarrow v_i$. And because $v''_j > v''_i > v'_i$ we have $v_i \not\rightarrow v_j$. Therefore $v_i \not\sim v_j$ contradicting that K is a clique.

It follows from (7) that the symbols $v'_1, v''_1, v'_2, v''_2, \dots, v'_k, v''_k$ must be arranged like this:

$$v'_{i_1} < v''_{i_1} < v'_{i_2} < v''_{i_2} < \dots < v'_{i_k} < v''_{i_k}.$$

The probability this happens—and this gives $E(X_K)$ —is $k!/(2k)!$.

Now let $X = \sum X_K$ where the sum is over all k -element subsets of B . This gives

$$\Pr(X > 0) \leq E(X) \leq \binom{n}{k} \frac{k!}{(2k)!} \leq \frac{n^k}{(2k)!}.$$

Put $k = 2\sqrt{n}$ and apply Stirling's formula to show that $E(X) \rightarrow 0$ as $n \rightarrow \infty$.

This implies that, with high probability, $\text{CL}(G[B]) \leq 2\sqrt{n}$ and the results follows. \square

Next we examine the independence number α of a random threshold digraph.

Proposition 34. *With high probability, the independence number of \mathbf{D}_n is asymptotic to $n/4$. Specifically, with high probability we have*

$$\frac{n}{4} - \omega(1)\sqrt{n} \leq \alpha(\mathbf{D}_n) \leq \frac{n}{4} + \omega(1)\sqrt{n \log n}.$$

Proof. In this proof we use the random representation model rather than the random order model. Let $X_1, Y_1, \dots, X_n, Y_n$ be iid uniform $[0, 1]$ random variables that produce the random threshold digraph \mathbf{D}_n .

First we verify the lower bound. Consider the set $I = \{v : X_v, Y_v < \frac{1}{2}\}$. Note that vertices in I form an independent set. The number of vertices in I is a binomial $(n, \frac{1}{4})$ random variable and so, with high probability, $\alpha[\mathbf{D}_n] \geq |I| \geq \frac{n}{4} - \omega(1)\sqrt{n}$.

Next we verify the upper bound. Let $U = \{v : X_v + Y_v \geq 1\}$. Check that for any two vertices, $v, w \in U$ we have $v \rightarrow w$ or $w \rightarrow v$, so a maximum independent set of \mathbf{D}_n can contain at most one vertex in U .

Let L be the complementary set of vertices, $\{v : X_v + Y_v < 1\}$ and let A be a largest independent subset of L . Note that $|A| \leq \alpha[\mathbf{D}_n] \leq |A| + 1$.

Among all the vertices in A , choose v with largest X value and w with largest Y value. It may be the case $v = w$. Let $t = X_v$.

We claim that $X_v + Y_w < 1$; here's why. If $v = w$, then $X_v + Y_v < 1$ because $v \in L$. If $v \neq w$ then $X_v + Y_w < 1$ because A is an independent set.

Therefore $Y_w < 1 - X_v = 1 - t$. It follows that for any vertex $a \in A$ we have $X_a \leq t$ and $Y_a < 1 - t$. That is, the points representing vertices in A lie in the rectangle $[0, t] \times [0, 1 - t]$.

For integer k with $1 \leq k \leq n$, define the rectangle R_k to be

$$R_k = \left[0, \frac{k}{n}\right] \times \left[0, 1 - \frac{k-1}{n}\right]$$

whose area is less than $\frac{1}{4} + \frac{1}{n}$. See Figure 10. Note that the rectangle $[0, t] \times [0, 1 - t]$ is

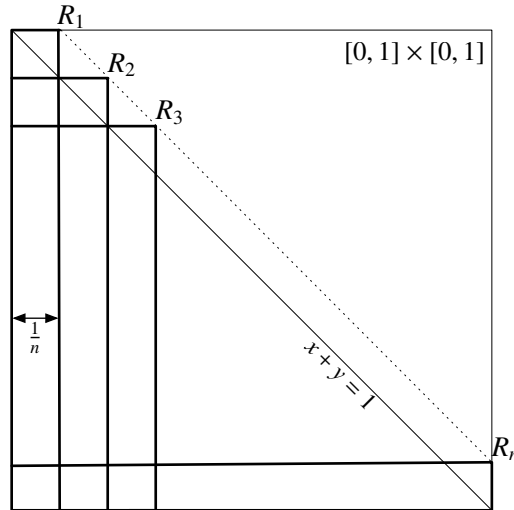


Figure 10: Covering the triangular region $\{(x, y) \in [0, 1]^2 : x + y \leq 1\}$ with rectangles R_1, R_2, \dots, R_n .

contained an R_k for some k .

The expected number of points (X_i, Y_i) in an R_k is less than $\frac{1}{4}n + 1$ so the probability a specific R_k contains more than $\frac{1}{4}n + 1 + \delta$ points is less than $\exp\{-c\delta^2/n\}$ for some absolute constant $c > 0$. Therefore, the probability *any* of the R_k contains more than $\frac{1}{4}n + 1 + \delta$ points is bounded by $n \exp\{-c\delta^2/n\}$. Taking $\delta \gg \sqrt{n \log n}$ is sufficient to conclude that, with high probability, $\alpha(\mathbf{D}_n) \leq \frac{n}{4} + 2 + \delta$. \square

Two comments on this result. First, the extra factor of $\sqrt{\log n}$ in the upper bound is likely just an artifact of our proof technique; we expect it can be eliminated. Second, define a *strong clique* in a digraph D to be a set of vertices K such that both $u \rightarrow v$ and $v \rightarrow u$ for all distinct $u, v \in K$. Then this result implies that the maximum size of a strong clique in \mathbf{D}_n is, with high probability, asymptotically $n/4$ (and we have the same bounds).

We now resume our analysis of the chromatic number of \mathbf{D}_n . We know that, with high probability, $\text{CL}[\text{simp}(\mathbf{D}_n)] \sim n/2$; this gives a lower bound on the chromatic number. In fact, the chromatic number equals the clique number.

Recall that a graph G is *perfect* if $\chi(H) = \text{CL}(H)$ for all induced subgraphs H of G . Berge [1] conjectured and Chudnovsky, Robertson, Seymour, and Thomas [3] proved the following landmark result.

Theorem 35 (Strong Perfect Graph Theorem). *A graph is perfect if and only if it does not contain, for any $k \geq 2$, an odd hole C_{2k+1} or an odd antihole $\overline{C_{2k+1}}$ as an induced subgraph.* \square

We employ this result to show that the underlying simple graph of a threshold digraph is perfect.

Theorem 36. *Let D be a threshold digraph and let $G = \text{simp}(D)$. Then G is perfect.*

We prove this by showing that G does not contain any odd holes or antiholes. That follows once we show that if $\text{simp}(D) = C_{2k+1}$ or $\text{simp}(D) = \overline{C_{2k+1}}$, then D is not a threshold digraph. A key step in doing this is the following result that characterizes those threshold digraphs D such that $\text{simp}(D) = P_4$.

Lemma 37. *Let D be a threshold digraph and suppose $\text{simp}(D) = P_4$ where the vertices are labeled such that we have $1 \sim 2 \sim 3 \sim 4$. Then the following conclusions hold:*

1. *We have both $1 \rightarrow 2$ and $3 \rightarrow 4$, or else we have both $2 \rightarrow 1$ and $4 \rightarrow 3$.*
2. *We cannot have both $1 \rightarrow 2$ and $4 \rightarrow 3$. Likewise, we cannot have both $2 \rightarrow 1$ and $3 \rightarrow 4$.*
3. *Consequently, we cannot have both $1 \rightarrow 2$ and $2 \rightarrow 1$. Nor can we have both $3 \rightarrow 4$ and $4 \rightarrow 3$.*

The conclusion to Lemma 37 is illustrated in Figure 11 in which we present all threshold digraphs D such that $\text{simp}(D) = P_4$ where $1 \sim 2 \sim 3 \sim 4$. Notice that the edge joining

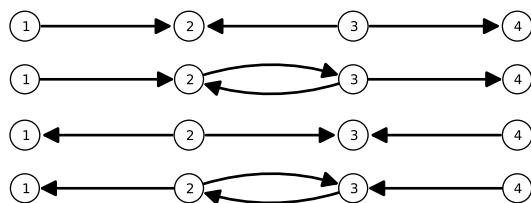


Figure 11: The threshold digraphs D with $\text{simp}(D) = P_4$ in which $1 \sim 2 \sim 3 \sim 4$.

1 and 2 and edge joining 3 and 4 must be oriented in the same direction, and in neither case do we have antiparallel edges.

Proof. Let D be any directed threshold graph such that $\text{simp}(D) = P_4$ with vertices labeled so that $1 \sim 2 \sim 3 \sim 4$.

We note that we cannot have both $1 \rightarrow 2$ and $4 \rightarrow 3$ in D because no matter how the edge $2 \sim 3$ is directed, the digraph would contain a two switch, violating Theorem 8. Likewise, D cannot contain both $2 \rightarrow 1$ and $3 \rightarrow 4$.

The conclusions now follow. □

Lemma 38. *There is no threshold digraph D such that $\text{simp}(D) = 2K_2$.*

Proof. If $\text{simp}(D) = 2K_2$ then D contains a two-switch and therefore is not a threshold digraph. □

Lemma 39. *For all $n \geq 5$, there is no threshold digraph D such that $\text{simp}(D) = C_n$.*

Note that both odd and even holes on five or more vertices are ruled out.

Proof. Let $n \geq 5$ and suppose, for contradiction, that D is a threshold digraph with $\text{simp}(D) = C_n$.

If $n \geq 6$, then C_n contains $2K_2$ as an induced subgraph and this leads to a contradiction of Lemma 38.

It remains to consider the case $\text{simp}(D) = C_5$ in which, say, we have $1 \sim 2 \sim 3 \sim 4 \sim 5 \sim 1$.

We claim that D does not contain a pair of antiparallel edges. Otherwise, suppose $1 \rightarrow 2$ and $2 \rightarrow 1$ in D . Form the digraph D' by deleting vertex 5 from D . Then $\text{simp}(D) = P_4$ but the antiparallel edges joining vertices 1 and 2 contradicting Lemma 37.

Without loss of generality we have $1 \rightarrow 2$ in D (and not $2 \rightarrow 1$). Considering vertices 1, 2, 3, 4, Lemma 37 implies that we must have $3 \rightarrow 4$. Considering vertices 3, 4, 5, 1, Lemma 37 implies that $5 \rightarrow 1$. Considering 5, 1, 2, 3 gives $2 \rightarrow 3$. Considering 2, 3, 4, 5 gives $4 \rightarrow 5$. Therefore D is a directed five-cycle consisting exactly of the edges $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1$ and none of the reverse arcs.

But then $1 \rightarrow 2$ and $3 \rightarrow 4$ form a two-switch. $\Rightarrow \Leftarrow$ □

Next we examine those threshold digraphs D such that $\text{simp}(D) = C_4$.

Lemma 40. *Let D be a threshold digraph with $\text{simp}(D) = C_4$. Then D is isomorphic to a threshold digraph with exactly these edges: $1 \rightarrow 2 \leftarrow 3 \rightarrow 4 \leftarrow 1$.*

Proof. Let D be a threshold digraph with $\text{simp}(D) = C_4$ in which $1 \sim 2 \sim 3 \sim 4 \sim 1$. Without loss of generality, $1 \rightarrow 2$ in D (we have not yet ruled out $2 \rightarrow 1$).

The digraph D must contain $3 \rightarrow 4$ or $3 \leftarrow 4$, but the latter can be ruled out for otherwise D would contain a two-switch. By the same reasoning, D cannot also contain $2 \rightarrow 1$ for otherwise $2 \rightarrow 1$ and $3 \rightarrow 4$ would form a two-switch. Thus we have $1 \rightarrow 2$ and $3 \rightarrow 4$, but neither reverse arc.

By the same reasoning applied to $2 \sim 3$ and $4 \sim 1$ we have exactly one of (a) $2 \rightarrow 3$ and $4 \rightarrow 1$ or (b) $3 \rightarrow 2$ and $1 \rightarrow 4$. However, we can rule out (a) because then $1 \rightarrow 2$ and $3 \rightarrow 4$ would form a two-switch. □

We note that the digraph D from the conclusion to Lemma 40 is, in fact, a threshold digraph. Here is a representation that gives exactly the edges $1 \rightarrow 2 \leftarrow 3 \rightarrow 4 \leftarrow 1$:

v		1		2		3		4
$f(v)$		0.8		0.2		0.8		0.2
$g(v)$		0.1		0.5		0.1		0.5

Finally, we consider antiholes on five or more vertices.

Lemma 41. *For all $n \geq 5$, there is no threshold digraph D such that $\text{simp}(D) = \overline{C_n}$.*

Proof. By Lemma 39, there is no threshold digraph D with $\text{simp}(D) = C_5$, and since $\overline{C_5} \cong C_5$, we may assume $n \geq 6$.

Suppose, for contradiction, D is a threshold digraph with $\text{simp}(D) = \overline{C_n}$. Without loss of generality, $V(G) = V(D) = [n]$.

Note that $3 \sim 1 \sim 4 \sim 2$ is an induced P_4 in G . By Lemma 37 we must exactly one of the following holds in D :

- $1 \rightarrow 3$ and $2 \rightarrow 4$ but $3 \not\rightarrow 1$ and $4 \not\rightarrow 2$, or
- $3 \rightarrow 1$ and $4 \rightarrow 2$ but $1 \not\rightarrow 3$ and $2 \not\rightarrow 4$.

Without loss of generality, we have the former: $1 \rightarrow 3$ and $2 \rightarrow 4$.

Next consider vertices $\{2, 3, 4, 5\}$ and observe that $4 \sim 2 \sim 5 \sim 3$ is an induced P_4 in G . Since $2 \rightarrow 4$ in D , we must also have $3 \rightarrow 5$.

Likewise, considering vertices $\{3, 4, 5, 6\}$, knowing that $3 \rightarrow 5$ in D implies $4 \rightarrow 6$.

Continuing in this fashion, in D we have all of the following

$$1 \rightarrow 3, \quad 2 \rightarrow 4, \quad 3 \rightarrow 5, \quad \dots, \quad (n-2) \rightarrow n, \quad (n-1) \rightarrow 1, \quad \text{and} \quad n \rightarrow 2$$

but none of the antiparallel arcs from $(k+2)$ to k (modulo n). See Figure 12.

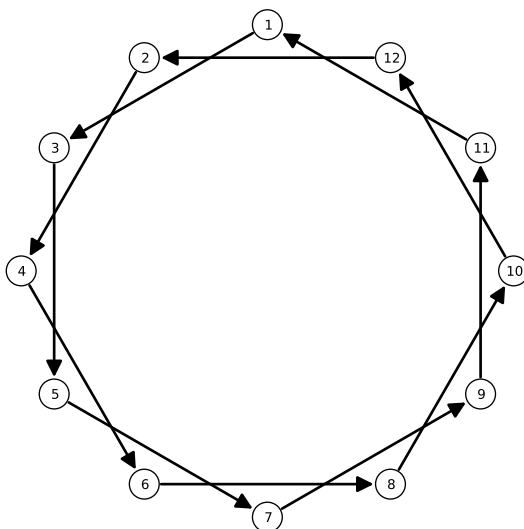


Figure 12: If a threshold digraph D satisfies $\text{simp}(D) = \overline{C_n}$, then arcs of the form $k \rightarrow (k+2)$ are oriented consistently.

We now consider vertices $\{2, 3, n-1, n\}$. These induce a C_4 in G : $2 \sim (n-1) \sim 3 \sim n \sim 2$. Applying Lemma 40, we know that the orientation of these edges in D must alternate. Since we have previously established that $n \rightarrow 2$ we conclude that $(n-1) \rightarrow 3$.

Next consider vertices $\{3, 4, n-2, n-1\}$. As before, these form a four-cycle: $3 \sim (n-2) \sim 4 \sim (n-1) \sim 3$. In D we have $(n-1) \rightarrow 3$ which implies $(n-2) \rightarrow 4$.

We continue this argument with vertices $\{4, 5, n-3, n-2\}$ and so forth and determine that D contains all arcs of the form $(n+2-k) \rightarrow k$, where $2 \leq k < n/2$ but none of their reversals. In other words, these arcs:

$$n \rightarrow 2, \quad (n-1) \rightarrow 3, \quad (n-2) \rightarrow 4, \quad \dots, \quad \left\lceil \frac{n}{2} + 2 \right\rceil \rightarrow \left\lfloor \frac{n}{2} \right\rfloor.$$

See Figure 13.

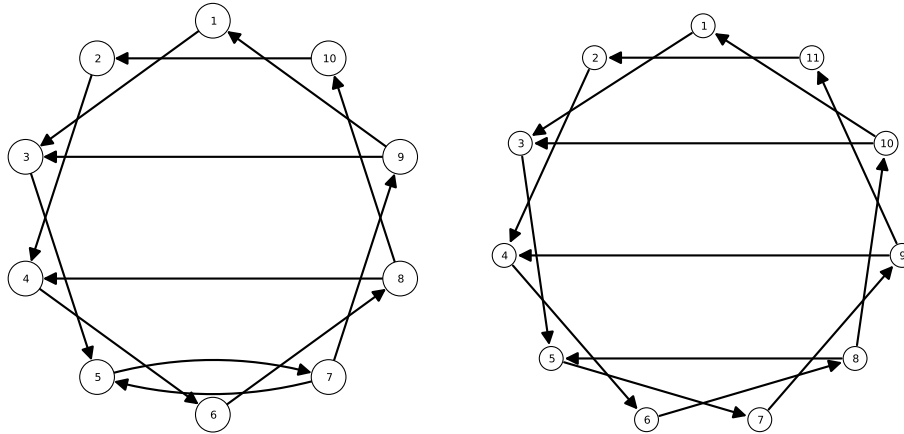


Figure 13: Additional arcs implied by $\text{simp}(D) = \overline{C_n}$.

If n is even, this implies that D contains both $(n/2) \rightarrow (n/2+2)$ and $(n/2+2) \rightarrow (n/2)$, a contradiction.

When n is odd, the consecutive vertices $(n-1)/2, (n+1)/2, (n+3)/2, (n+5)/2$ induce a P_4 in G that are oriented like this in D :

$$\frac{n+1}{2} \rightarrow \frac{n+5}{2} \rightarrow \frac{n-1}{2} \rightarrow \frac{n+3}{2}.$$

(See the path $6 \rightarrow 8 \rightarrow 5 \rightarrow 7$ in the right digraph in Figure 13.) This contradicts Lemma 37.

Therefore, for $n \geq 5$, there is no threshold digraph D with $\text{simp}(D) = \overline{C_n}$. □

Proof of Theorem 36. This follows directly from Theorem 36 by appealing to Lemmas 39 and 41. □

Corollary 42. *With high probability, the chromatic number of \mathbf{D}_n is asymptotic to $n/2$. Specifically, with high probability, we have*

$$\frac{n}{2} - \omega(1)\sqrt{n} \leq \chi(\mathbf{D}_n) \leq \frac{n}{2} + \omega(1)\sqrt{n}. \quad \square$$

References

- [1] Claude Berge. Färbung von Graphen, deren sämtliche bzw. deren ungerade Kreise starr sind. *Wiss. Z. Martin-Luther-Univ. Halle-Wittenberg Math.-Natur. Reihe*, 10:114, 1961.
- [2] Graham Brightwell and Peter Winkler. Counting linear extensions is #P-complete. In *Proceedings of the Twenty-third Annual ACM Symposium on Theory of Computing, STOC '91*, pages 175–181, New York, NY, USA, 1991. ACM.

- [3] Maria Chudnovsky, Neil Robertson, Paul Seymour, and Robin Thomas. The strong perfect graph theorem. *Annals of Mathematics*, 164:51–229, 2006.
- [4] V. Chvatal and P.L. Hammer. Aggregation of inequalities in integer programming. In P.L. Hammer, E.L. Johnson, B.H. Korte, and G.L. Nemhauser, editors, *Studies in Integer Programming*, volume 1 of *Annals of Discrete Mathematics*, pages 145–162. North-Holland, 1977.
- [5] Brian Cloteaux, M. Drew LaMar, Elizabeth Moseman, and James Shook. Threshold digraphs. Preprint, 2012. [arXiv:1212.1149](https://arxiv.org/abs/1212.1149)
- [6] Olivier Cogis. Ferrers digraphs and threshold graphs. *Discrete Mathematics*, 38(1):33–46, 1982.
- [7] Persi Diaconis, Susan Holmes, and Svante Janson. Threshold graph limits and random threshold graphs. *Internet Mathematics*, 5(3):267–320, 2008.
- [8] Martin C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Academic Press, 1980.
- [9] W. Hoeffding. A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19:293–325, 1948.
- [10] N.V.R. Mahadev and U.N. Peled. *Threshold Graphs and Related Topics*. North-Holland, 1995.
- [11] Elizabeth Reilly. *Random Threshold Graphs and Related Topics*. PhD thesis, Johns Hopkins University, 2009.
- [12] Elizabeth Reilly and Edward Scheinerman. Random threshold graphs. *Electronic Journal of Combinatorics*, 16(1), 2009. R130.
- [13] Christopher Ross. Properties of random difference graphs. *Electronic Journal of Combinatorics*, 19(4), 2012. P28.
- [14] Robert J. Serfling. *Approximation Theorems of Mathematical Statistics*. Wiley, 1980.
- [15] William T. Trotter. *Combinatorics and Partially Ordered Sets: Dimension Theory*. Johns Hopkins University Press, 1992.
- [16] Stephen J. Young and Edward Scheinerman. Directed random dot product graphs. *Internet Mathematics*, 5(1–2):91–112, 2008.