On small subgraphs in a random intersection digraph

Valentas Kurauskas*
Vilnius University, Naugarduko 24, LT-03225 Vilnius, Lithuania
Institute of Mathematics and Informatics, Akademijos 4, LT-08663 Vilnius, Lithuania
e-mail: valentas@gmail.com

26 October 2012

Abstract

Given a set of vertices V and a set of attributes W let each vertex $v \in V$ include an attribute $w \in W$ into a set $S^-(v)$ with probability p_- and let it include w into a set $S^+(v)$ with probability p_+ independently for each $w \in W$. The random binomial intersection digraph on the vertex set V is defined as follows: for each $u, v \in V$ the arc uv is present if $S^-(u)$ and $S^+(v)$ are not disjoint. For any $h = 2, 3, \dots$ we determine the birth threshold of the complete digraph on h vertices and describe the configurations of intersecting sets that realise the threshold.

Keywords: digraph, clique, threshold, random intersection graph.

1 Introduction

In the random intersection graph introduced by Karoński, Scheinerman and Singer-Cohen [13] (see also Godehardt and Jaworski [10]) we have a set of vertices V of size n and an additional set W of attributes (also sometimes called properties or keys) of size m. Each vertex v of V chooses a random subset of attributes in W and an edge uv is added to the graph if and only if the subsets of u and v intersect.

*The work was supported in part by the Research Council of Lithuania grant MIP-52/2010.

We consider a directed random intersection graph D on the vertex set V defined as follows (see [4]). Let each vertex v choose not one, but two random subsets: an "in"-subset, $S^+(v)$, and an "out"-subset, $S^-(v)$. An arc from vertex u to v is inserted in D whenever $S^-(u)$ intersects $S^+(v)$. Assuming, in addition, that each attribute $w \in W$ is included in the subset $S^-(v)$ with probability p_- and in the subset $S^+(v)$ with probability p_+ independently and independently of all other inclusions, we obtain a *random binomial intersection digraph* denoted $D = D(n, m, p_-, p_+)$.

In [18] a network of co-authors of mathematical papers is mentioned as an illustration for random intersection graphs. One might alternatively define a *citation digraph* where V is a set of mathematicians and we draw an arc from u to v if and only if u has cited v. The underlying set W here would be the set of all mathematical papers; and $S^-(u)$ (respectively, $S^+(u)$) would correspond to the set of papers u has cited (respectively, co-authored).

The random intersection graph model has received a lot of attention recently due to several different applications [8, 10, 13]. Properties such as thresholds for small graphs [13], degree distribution [3, 6, 18], formation of the giant component [1, 5], connectivity [2, 15] and clustering [6] have been studied, see also [17, 9, 14]. In some applications considering directed intersection graphs makes sense and might lead to more precise/adequate models. In particular, one may obtain a digraph with power law indegree distribution and bounded outdegree distribution. In addition these digraphs have a clustering property when m is of order n [4].

In the problem of determining the birth threshold of small subgraphs one is interested in the question of how dense a graph should be to have a desired subgraph with certainty. There is a rich literature devoted to birth thresholds in random graphs with independent edges where each edge appears with the same probability, see, e.g., Chapter 3 of [11]. The threshold for a random (binomial) intersection graph to contain a fixed subgraph has been studied in [13].

Here we consider a similar problem for random intersection digraphs. Let $\overrightarrow{K}_h$ be the complete digraph on vertex set $[h] = \{1, \ldots, h\}$ containing arcs xy and yx for each pair of distinct vertices $x, y \in [h]$. We aim to determine critical values of the parameters for $D(n, m, p_-, p_+)$ to have with a high probability a subgraph isomorphic to $\overrightarrow{K}_h$.

Given two finite sets C^- and C^+ we consider the ordered pair $C = (C^-; C^+)$ and the digraph $D(C)$ on the vertex set $C^- \cup C^+$ with the set of arcs $\{uv : u \in C^-, v \in C^+\}$. We call the pair C a *diclique*. We say that C is *proper* if C^-, C^+ are non-empty, otherwise say that it is *improper*. We

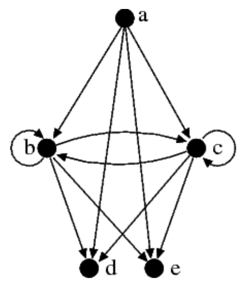

Figure 1: The diclique $(\{a,b,c\};\{b,c,d,e\})$.

remark that if the digraph $D(C)$ is non-empty then C must be proper and $D(C) \neq D(C')$ for $C \neq C'$. Therefore we will identify a proper diclique C with the corresponding digraph $D(C)$, see Figure 1.

To our knowledge, the diclique digraphs were first studied by Haralick [12], but in a different context.

In the random digraph D with vertex set V and attribute set W each attribute $w \in W$ defines a diclique $C(w) = (C^-(w); C^+(w))$ given by $C^-(w) = \{v \in V : w \in S^-(v)\}$ and $C^+(w) = \{v \in V : w \in S^+(v)\}$. It is convenient to interpret each attribute $w \in W$ as a distinct colour. Then all the attributes in W give rise to a family of dicliques of different colours which covers all arcs of D.

The paper is organised as follows. In the next section we present our main results. In Section 3 we give a general lemma for the birth threshold of a fixed directed graph H. In Section 4 we study a few special diclique covers of $\overrightarrow{K}_h$ and prove our main results Theorem 2.1 and Theorem 2.2.

We remind some standard notation used in the paper. For functions $f, g : \mathbb{N} \to \mathbb{R}_+$ we write $f \sim g$ if $\lim_{k\to\infty} f(k)/g(k) = 1$. We write $f = O(g)$ if $\limsup_{k\to\infty} f(k)/g(k) < \infty$, $f = \Omega(g)$ if $g = O(f)$ and $f = \Theta(g)$ if both $f = O(g)$ and $g = O(f)$. We write $f = o(g)$ if $f(k)/g(k) \to 0$.

Finally, thanks to an anonymous reviewer the author became aware of a related and very relevant result on the Poisson approximation of the number of cliques in sparse random intersection graphs by Rybarczyk and Stark [16].

2 Results

Before stating our main results we need to introduce some definitions related to dicliques. Without loss of generality we will assume that the set of vertices of the random digraph D is $V = [n]$.

For any diclique C, we call $V(C) = C^- \cup C^+$ the vertex set of C. Let $\mathcal{C} = \{C_1, C_2, \dots, C_s\}$ be a family of dicliques, (we allow $\mathcal{C}$ to be a multiset and in this paper we consider only finite families $\mathcal{C}$). Let us denote by $V(\mathcal{C})$ the union of all vertices of the dicliques, $V(\mathcal{C}) = \bigcup V(C_i)$. We say that D *contains* $\mathcal{C}$ if there are distinct attributes $w_1, \dots, w_s \in W$, such that $C_i \subseteq C(w_i)$ for each $i = 1, \dots, s$ (the set operations for dicliques are defined componentwise). Also, let us call a diclique family *proper* if all its dicliques are proper.

Let $\mathcal{C}$ be any diclique family with $V(\mathcal{C}) = \{v_1, \dots, v_r\} \subseteq [n]$ and assume that $v_1 < \cdots < v_r$. For any set $S = \{x_1, \dots, x_r\} \subseteq [n]$ with $x_1 < \cdots < x_r$, let us denote by $M(\mathcal{C}, S)$ the diclique family which is an image of $\mathcal{C}$ obtained by renaming v_i to x_i for each $i = 1, \dots, r$. We call $M(\mathcal{C}, S)$ a *copy* of $\mathcal{C}$.

Each family of dicliques $\mathcal{C}$ defines a digraph $H = H(\mathcal{C})$ with vertex set $V(\mathcal{C})$: an arc is present in H whenever it is present in some $D(C)$, $C \in \mathcal{C}$. We say that the family $\mathcal{C}$ is a *diclique cover* of H.

The digraph $\overrightarrow{K}_h$ can be covered by dicliques in many different ways. Consider the following important symmetric diclique covers of $\overrightarrow{K}_h$:

- $\mathcal{C}_M = \{([h]; [h])\}$, the monochromatic diclique cover;
- $\mathcal{C}_R = E(\overrightarrow{K}_h)$, the rainbow diclique cover, where $E(\overrightarrow{K}_h)$ is the set of arcs of $\overrightarrow{K}_h$ and we identify each arc uv with the diclique $(\{u\}; \{v\})$;
- $\mathcal{C}_{in} = \{([h] \setminus \{i\}; \{i\}) \,:\, i \in [h]\}$, the cover by in-stars;
- $\mathcal{C}_{out} = \{(\{i\}; [h] \setminus \{i\}) \,:\, i \in [h]\}$, the cover by out-stars.

The motivation for the names "monochromatic" and "rainbow" is that a single attribute (or colour) $w \in W$ may generate a copy of $\mathcal{C}_M$ in D, while $h(h-1)$ attributes are needed for a copy of $\mathcal{C}_R$.

We will consider a sequence of random digraphs $\{D(k), k = 1, 2, \dots\}$ where $D(k) = D(n, m, p_-, p_+)$, $n = n(k)$ (we always assume that $n(k)$ is increasing), $m = m(k)$, $p_- = p_-(k)$ and $p_+ = p_+(k)$ all depend on k. If not stated otherwise all limits below are as $k \to \infty$.

Let us now define what a birth threshold function for $\overrightarrow{K}_h$ is. We would like to have a function that, for a sequence of random digraphs $\{D(k)\}$, indicates whether a copy of $\overrightarrow{K}_h$ is present whp. Since the sequence $\{D(k)\}$ depends on several parameters which are themselves sequences, such a function has to take into account all of them.

Let X be either a fixed digraph or a diclique family. Let $X \in D$ denote the event that the random digraph D contains a copy of X. Given a sequence of random digraphs $\{D(k)\}$ we call a function $\tau : \mathbb{N}^2 \times [0;1]^2 \to \mathbb{R}_+$ a *birth threshold function* for X if both of the following implications hold:

$$\tau(n,m,p_-,p_+) \to 0 \implies P(X \in D(k)) \to 0;$$
$$\tau(n,m,p_-,p_+) \to \infty \implies P(X \in D(k)) \to 1.$$

Given a sequence of random digraphs $\{D(k)\}$ and a birth threshold function τ for $\overrightarrow{K}_h$ we call a collection $\mathcal{L}$ of diclique covers of $\overrightarrow{K}_h$ the *leading set* if τ is a birth threshold function for each $\mathcal{C} \in \mathcal{L}$ and the following implications hold:

1) $\tau(n,m,p_-,p_+) = O(1) \implies P(\mathcal{C}' \in D(k)) \to 0$ for each proper diclique cover $\mathcal{C}'$ of $\overrightarrow{K}_h$ such that $\mathcal{C}' \notin \mathcal{L}$;

2) $\tau(n,m,p_-,p_+) = \Theta(1) \implies P(\mathcal{C} \in D(k)) = \Omega(1)$ for each $\mathcal{C} \in \mathcal{L}$.

In the case where $\mathcal{L}$ consists just of a single diclique family $\mathcal{C}$, we call $\mathcal{C}$ the *leading cover*.

We will consider the following birth threshold functions:

$$\tau_1 = nm^{1/h}p_-p_+; \qquad \tau_2 = n^{1/(h-1)}mp_-p_+;$$
$$\tau_3 = nmp_-^{h-1}p_+; \qquad \tau_4 = nmp_-p_+^{h-1}.$$

We are now ready to state our main result (see Figure 2 for an illustration).

Theorem 2.1 *Let $h \geq 3$ be a fixed integer. Write $\alpha_0 = 1 - \frac{1}{(h-1)^2}$. Let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that n is increasing, $m = \Theta(n^\alpha)$ for some $\alpha > 0$, $p_- \to 0$ and $p_+ \to 0$.*

(i) If $\alpha < \alpha_0$ and

- *(a) $m^{\frac{h-1}{h(h-2)}}p_- \to \infty$ then τ_3 is a birth threshold function for $\overrightarrow{K}_h$ with the leading cover $\mathcal{C}_{in}$;*
- *(b) $m^{\frac{h-1}{h(h-2)}}p_- \to 0$ and $m^{\frac{h-1}{h(h-2)}}p_+ \to 0$ then τ_1 is a birth threshold function for $\overrightarrow{K}_h$ with the leading cover $\mathcal{C}_M$;*
- *(c) $m^{\frac{h-1}{h(h-2)}}p_+ \to \infty$ then τ_4 is a birth threshold function for $\overrightarrow{K}_h$ with the leading cover $\mathcal{C}_{out}$.*

(ii) If $\alpha \geq \alpha_0$ and

(a) $mp_+ \to 0$ *then* τ_3 *is a birth threshold function for* $\overrightarrow{K}_h$ *with the leading cover* $\mathcal{C}_{in}$*;*

(b) $\alpha \neq \alpha_0$*,* $mp_- \to \infty$ *and* $mp_+ \to \infty$ *then* τ_2 *is a birth threshold function for* $\overrightarrow{K}_h$ *with the leading cover* $\mathcal{C}_R$*;*

(c) $mp_- \to 0$ *then* τ_4 *is a birth threshold function for* $\overrightarrow{K}_h$ *with the leading cover* $\mathcal{C}_{out}$*.*

Let us introduce the following two collections of diclique covers. Let $\mathcal{S}_{in}$ be the collection of all proper diclique covers of $\overrightarrow{K}_h$ such that each $\mathcal{C} \in \mathcal{S}_{in}$ is a set of arc-disjoint in-stars, that is, for each $C \in \mathcal{C}$ we have $|C^+| = 1$ and for each $C_i, C_j \in \mathcal{C}$ with $i \neq j$ and $C_i^+ = C_j^+$ we have $C_i^- \cap C_j^- = \emptyset$. Similarly, let $\mathcal{S}_{out}$ be the collection of all proper diclique covers $\mathcal{C}$ of $\overrightarrow{K}_h$ such that for each $C \in \mathcal{C}$ we have $|C^-| = 1$ and for each $C_i, C_j \in \mathcal{C}$ with $i \neq j$ and $C_i^- = C_j^-$ we have $C_i^+ \cap C_j^+ = \emptyset$. Observe that $\mathcal{C}_R, \mathcal{C}_{in} \in \mathcal{S}_{in}$ and $\mathcal{C}_R, \mathcal{C}_{out} \in \mathcal{S}_{out}$.

For the "boundary" cases of the parameters in Theorem 2.1 we have:

Theorem 2.2 *Let* $h, \{D(k)\}, n, \alpha_0, m$ *be as in Theorem 2.1. Suppose* $p_- \to 0$ *and* $p_+ \to 0$*. If*

(a) $\alpha < \alpha_0$ *and* $m^{\frac{h-1}{h(h-2)}} p_- = \Theta(1)$ *then* τ_1, τ_3 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\{\mathcal{C}_M, \mathcal{C}_{in}\}$*.*

(b) $\alpha < \alpha_0$ *and* $m^{\frac{h-1}{h(h-2)}} p_+ = \Theta(1)$ *then* τ_1, τ_4 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\{\mathcal{C}_M, \mathcal{C}_{out}\}$*.*

(c) $\alpha = \alpha_0$*,* $mp_- \to \infty$ *and* $mp_+ \to \infty$ *then* τ_1, τ_2 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\{\mathcal{C}_M, \mathcal{C}_R\}$*.*

(d) $\alpha > \alpha_0$ *and* $mp_+ = \Theta(1)$ *then* τ_2, τ_3 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\mathcal{S}_{in}$*.*

(e) $\alpha > \alpha_0$ *and* $mp_- = \Theta(1)$ *then* τ_2, τ_4 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\mathcal{S}_{out}$*.*

(f) $\alpha = \alpha_0$ *and* $mp_+ = \Theta(1)$ *then* τ_1, τ_2, τ_3 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\{\mathcal{C}_M\} \cup \mathcal{S}_{in}$*.*

(g) $\alpha = \alpha_0$ *and* $mp_- = \Theta(1)$ *then* τ_1, τ_2, τ_4 *are birth threshold functions for* $\overrightarrow{K}_h$ *with the leading set* $\{\mathcal{C}_M\} \cup \mathcal{S}_{out}$*.*

We note that the argument used in the proof of Theorem 2.1 can also be extended to the case where p_- or p_+ are bounded away from zero. In this case the birth threshold function remains the same, but the leading sets have to be slightly modified.

Finally, for the case $h = 2$ we have

Remark 2.3 *Let $\{D(k)\}, n, m$ be as in Theorem 2.1. Suppose $h = 2, p_- \to 0$ and $p_+ \to 0$. Then τ_2 is a birth threshold function for $\overrightarrow{K}_2$ with the leading cover $\mathcal{C}_R$.*

3 Diclique covers: a general lemma

In this section we present some important estimates and Lemma 3.2 that relates the birth threshold of a fixed digraph with presence of its diclique covers. This is very similar to results for undirected random intersection graphs, see Theorem 3 of [13]. We postpone the proofs of the estimates (1) - (8) of this section till Appendix A.

Let $\{D(k)\}$ be a sequence of random binomial intersection digraphs and write $D = D(k)$. Recall that we denote the vertex set of D by V and the attribute set by W. Suppose $m = \Theta(n^\alpha)$ for some $\alpha > 0$. In (1) - (5) we will assume that

$$p_- \to 0, \quad p_+ \to 0 \quad \text{and} \quad mp_-p_+ \to 0.$$

Note that in this case mp_-p_+ is asymptotically equivalent to the probability that a fixed directed edge exists.

Given a diclique $C = (C^-; C^+)$ and a set S the restriction of C to S is the diclique $C[S] = (C^- \cap S; C^+ \cap S)$. The restriction for a diclique family is defined by $\mathcal{C}[S] = \{C[S] : C \in \mathcal{C}, V(C) \cap S \neq \emptyset\}$.

Let $S \subseteq V$ and let C be a diclique with $V(C) \subseteq S$. We will say that a *monochromatic C occurs on S* if in the realization of D there is at least one attribute $w \in W$ such that $C = C(w)[S]$ (we say that w *generates* C on S). We denote the probability of the event that a monochromatic C occurs on S by $P(C)$.

We say that a diclique family $\mathcal{C} = \{C_1, \ldots, C_s\}$ is *induced* in D if there are distinct attributes $w_1, \ldots, w_s \in W$ such that $C_i = C(w_i)[V(\mathcal{C})]$, for each $i = 1, \ldots, s$ and for any attribute $w \in W \setminus \{w_1, \ldots, w_s\}$ the diclique $C(w)[V(\mathcal{C})]$ is improper. Thus if $\mathcal{C}$ is induced in D then D contains $\mathcal{C}$ (see Section 2). We denote the probability of the event that $\mathcal{C}$ is induced in D by $P(\mathcal{C})$. Let $S_1, S_2, \ldots, S_N$ be all the subsets of V of size $r = |V(\mathcal{C})|$, where

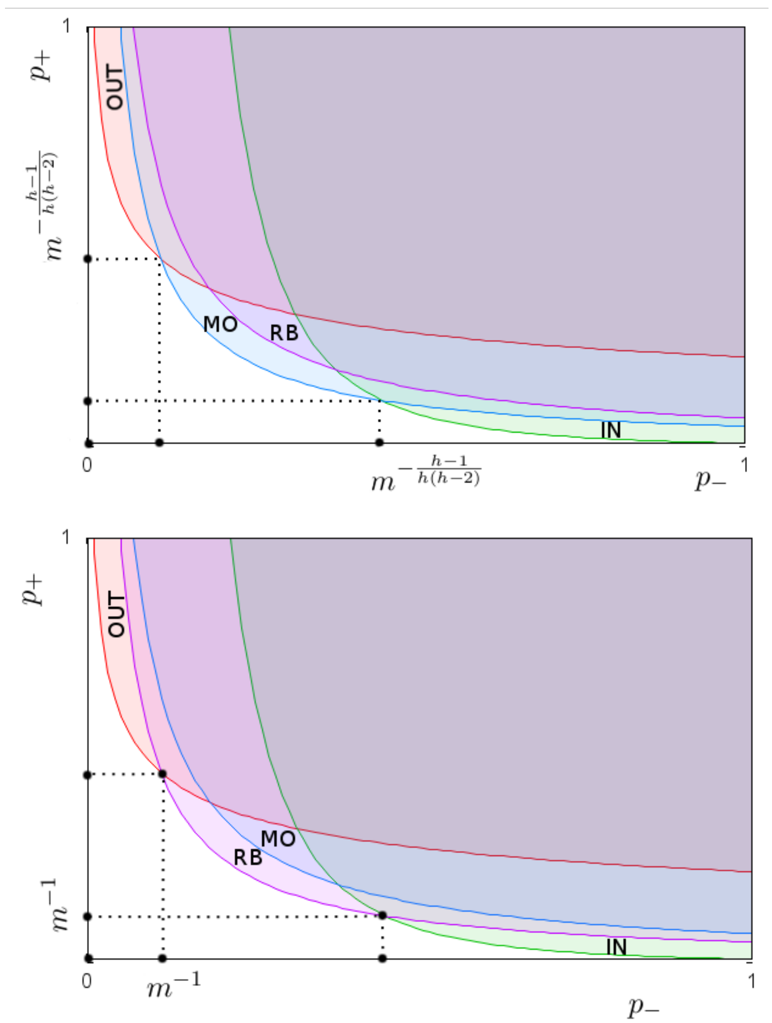

Figure 2: Schematic illustration of Theorems 2.1 and 2.2. Top: $\alpha < \alpha_0$, bottom: $\alpha > \alpha_0$. The coloured area is the region where D contains a copy of the special diclique cover whp (IN = "in-stars" cover $\mathcal{C}_{in}$, OUT="out-stars" cover $\mathcal{C}_{out}$, MO=monochromatic cover $\mathcal{C}_M$, RB = "rainbow" cover $\mathcal{C}_R$), the remaining area is where D does not contain that diclique cover whp. The white area is where D does not have a copy of $\overrightarrow{K}_h$ whp. On the contour between the white and the coloured area we know that D does not contain any other proper diclique cover whp (excluding the black points that correspond to the collections $\mathcal{S}_{in}$ and $\mathcal{S}_{out}$).

$N = \binom{n}{r}$. Let $\mathbb{I}_i = \mathbb{I}_{M(\mathcal{C},S_i)}$ be the indicator of the event that the copy $M(\mathcal{C}, S_i)$ is induced in D. Then the *number of induced copies* of $\mathcal{C}$ in D is defined by

$$X(\mathcal{C}) = \sum_{i=1}^{N} \mathbb{I}_i.$$

Let $S \subseteq V$ and suppose $C = (C^-; C^+)$ is a proper diclique such that $V(C) \subseteq S$. Then the probability that a monochromatic C occurs on S is

$$P(C) \sim \tilde{P}(C) := mp_-^{|C^-|} p_+^{|C^+|}. \tag{1}$$

Following [13], for a diclique family $\mathcal{C} = \{C_i\}$ we write $\sum \mathcal{C}^+ = \sum |C_i^+|$ and $\sum \mathcal{C}^- = \sum |C_i^-|$, and we denote by $|\mathcal{C}|$ the cardinality of $\mathcal{C}$.

For a diclique family $\mathcal{C}$ let $C_1, \ldots, C_t$ be all its *distinct* dicliques and let $a_1, \ldots, a_t$ be their multiplicities in $\mathcal{C}$. Let us denote $a_{\mathcal{C}} = (a_1! a_2! \ldots a_t!)^{-1}$.

Fix a diclique family $\mathcal{C}$ with $V(\mathcal{C}) \subseteq V$. If $\mathcal{C}$ is proper then the probability that $\mathcal{C}$ is induced in D is

$$P(\mathcal{C}) \sim \tilde{P}(\mathcal{C}) := a_{\mathcal{C}} \prod_{C \in \mathcal{C}} \tilde{P}(C) = a_{\mathcal{C}} m^{|\mathcal{C}|} p_-^{\sum \mathcal{C}^-} p_+^{\sum \mathcal{C}^+}. \tag{2}$$

Write

$$\mu(\mathcal{C}) = \mu(\mathcal{C}, n, m, p_-, p_+) := \frac{n^{|V(\mathcal{C})|}}{|V(\mathcal{C})|!} \tilde{P}(\mathcal{C}) = \frac{a_{\mathcal{C}}}{|V(\mathcal{C})|!} n^{|V(\mathcal{C})|} m^{|\mathcal{C}|} p_-^{\sum \mathcal{C}^-} p_+^{\sum \mathcal{C}^+}.$$

If $\mathcal{C}$ is proper then the number $X(\mathcal{C})$ of induced copies of $\mathcal{C}$ in D satisfies

$$\mathbb{E}\, X(\mathcal{C}) \sim \mu(\mathcal{C}). \tag{3}$$

More generally, suppose $\mathcal{C}$ is not necessarily proper. Suppose that, in addition, the following technical assumption is satisfied: for each $j = 0, \ldots, |V(\mathcal{C})|$

$$mp_-^j \to a_j \quad \text{and} \quad mp_+^j \to b_j \tag{4}$$

for some $a_j, b_j \in [0; \infty]$. Let $\hat{\mathcal{C}}$ be the diclique family obtained from $\mathcal{C}$ by taking only those dicliques $C \in \mathcal{C}$ that satisfy $mp_-^{|C^-|} p_+^{|C^+|} \to 0$ (for proper $\mathcal{C}$ we always have $\hat{\mathcal{C}} = \mathcal{C}$). Define

$$\tilde{\mu}(\mathcal{C}) = \tilde{\mu}(\mathcal{C}, n, m, p_-, p_+) := n^{|V(\mathcal{C})|} m^{|\hat{\mathcal{C}}|} p_-^{\sum \hat{\mathcal{C}}^-} p_+^{\sum \hat{\mathcal{C}}^+}.$$

Then

$$\mathbb{E}\, X(\mathcal{C}) = \Theta\left(\tilde{\mu}(\mathcal{C})\right). \tag{5}$$

Observe that the event "$\mathcal{C}$ is induced in D" allows any number of extra improper dicliques on $V(\mathcal{C})$.

For a proper diclique family $\mathcal{C}$ and any sequence of random graphs $\{D(k)\}$ (without any assumptions on p_-, p_+, m) we have

$$P(D(k) \text{ contains } \mathcal{C}) \leq \tilde{P}(\mathcal{C}); \tag{6}$$
$$P(\mathcal{C} \in D(k)) \leq \mu(\mathcal{C}); \tag{7}$$

(recall that the first event concerns a fixed diclique cover while the second one asks for any copy of $\mathcal{C}$ in $D(k)$) and, if (4) holds then for any diclique family $\mathcal{C}$

$$P(\mathcal{C} \in D(k)) \leq \tilde{\mu}(\mathcal{C}). \tag{8}$$

We will use the following simple technical lemma several times below.

Lemma 3.1 *For a positive integer t let $f_1, \ldots, f_t, g : \mathbb{N} \to \mathbb{R}$ be any functions. Suppose that for any increasing sequence of positive integers (n_k), $k = 1, 2 \ldots$ such that*

$$\lim_{k\to\infty} f_i(n_k) \text{ exists or is in } \{-\infty, \infty\} \text{ for each } i = 1, \ldots, t \tag{9}$$

we have $\lim_{k\to\infty} g(n_k) = b \in [-\infty; \infty]$. Then $\lim_{n\to\infty} g(n) = b$.

Proof Write $\bar{g} = \limsup_{n\to\infty} g(n)$. Then there is an increasing sequence of integers $(n_j), j = 1, 2, \ldots$ such that $g(n_j) \to \bar{g}$. By the Weierstrass theorem this sequence has a subsequence $(n_k), k = 1, 2, \ldots$ such that (9) holds. So $b = \lim_{k\to\infty} g(n_k) = \bar{g}$. Similarly $b = \liminf_{n\to\infty} g(n)$ and the claim follows. $\square$

Let us call a diclique family $\mathcal{C}$ *simple* if it is proper and has no repetitive elements. The following result allows to find a birth threshold of a fixed digraph by considering just a constant number (which depends on h) of diclique covers.

Lemma 3.2 *Let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that n is increasing. Let $h \geq 2$ be an integer, and let H be a digraph with $V(H) = [h]$ and without isolated vertices. Suppose that $p_- \to 0$, $p_+ \to 0$ and $mp_-p_+ \to 0$. Then*

(a) if for each simple diclique cover $\mathcal{C}$ of H there is a non-empty set $S \subseteq V(\mathcal{C})$ such that $\mathbb{E}\, X(\mathcal{C}[S]) \to 0$ then whp $D(k)$ does not contain a copy of H;

(b) *if there is a simple diclique cover $\mathcal{C}$ of H such that $\mathbb{E}\, X(\mathcal{C}[S]) \to \infty$ for each non-empty set $S \subseteq V(\mathcal{C})$ then whp $D(k)$ contains an induced copy of $\mathcal{C}$ (and therefore also a copy of H as an induced subgraph).*

Proof The part (a) is easy. Suppose the sequence $\{D(k)\}$ satisfies the conditions of (a), but $\limsup P(\overrightarrow{K}_h \in D) > 0$, where $D = D(k)$. Since the event $\overrightarrow{K}_h \in D$ implies $\mathcal{C} \in D$ for some simple diclique cover $\mathcal{C}$ of $\overrightarrow{K}_h$ and the number of such diclique covers is finite we have $\limsup P(\mathcal{C}_0 \in D) > 0$ for one of such covers $\mathcal{C}_0$. By Lemma 3.1 we may assume that (4) holds. Take a set $S \subseteq [h]$ such that $\mathbb{E}\,(X(\mathcal{C}_0[S])) \to 0$. By (5) and (8) we have

$$P(\mathcal{C}_0 \in D) \leq P(\mathcal{C}_0[S] \in D) \leq \tilde{\mu}(\mathcal{C}_0[S]) = \Theta(\mathbb{E}\, X(\mathcal{C}_0[S])) = o(1)$$

which is a contradiction.

Now let us prove (b). Let $\mathcal{C}$ be a simple diclique cover of H that satisfies the condition in (b). Note that $|V(\mathcal{C})| = h$ since H has no isolated vertices. Recall that the number of induced copies of $\mathcal{C}$ in D is $X = X(\mathcal{C}) = \sum_{i=1}^{N} \mathbb{I}_i$ where $\mathbb{I}_i = 1$ if and only if the diclique cover $M(\mathcal{C}, S_i)$ is induced in D and by (3)

$$\mathbb{E}\, X \sim \mu(\mathcal{C}) \sim \frac{n^h}{h!} \tilde{P}(\mathcal{C}) \to \infty.$$

So the claim will follow by the method of second moments if we show that $Var(X)/(\mathbb{E}\, X)^2 \to 0$. We have

$$\mathbb{E}\, X^2 = \mathbb{E}\, X + \sum_{i \neq j} \mathbb{I}_i \mathbb{I}_j.$$

If the sets S_i and S_j do not intersect, we have that $\mathbb{I}_i$ and $\mathbb{I}_j$ are independent and $\mathbb{E}\, \mathbb{I}_i \mathbb{I}_j = \mathbb{E}\, \mathbb{I}_i \mathbb{E}\, \mathbb{I}_j$. There are in total $\binom{n}{h}\binom{n-h}{h}$ (ordered) pairs of sets that do not intersect. So

$$\begin{aligned} Var(X) = \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2 \sim \mathbb{E}\, X + \frac{n^{2h}}{(h!)^2}\,(1 + o(1))\, P(\mathcal{C})^2 \\ - \frac{n^{2h}}{(h!)^2}\,(1 + o(1))\, P(\mathcal{C})^2 + \sum_{S_i \cap S_j \neq \emptyset} \mathbb{E}\, \mathbb{I}_i \mathbb{I}_j \\ = o\left((\mathbb{E}\, X)^2\right) + \sum_{S_i \cap S_j \neq \emptyset} \mathbb{E}\, \mathbb{I}_i \mathbb{I}_j; \end{aligned}$$

and we have that

$$\frac{\sum_{S_i \cap S_j \neq \emptyset} \mathbb{E}\, \mathbb{I}_i \mathbb{I}_j}{(\mathbb{E}\, X)^2} \leq \frac{\sum_{s=1}^{h} \binom{n}{h}\binom{h}{s}\binom{n-h}{h-s} \max_{|S_i \cap S_j| = s} \mathbb{E}\, \mathbb{I}_i \mathbb{I}_j}{(\mathbb{E}\, X)^2} \leq C \sum_{s=1}^{h} T_s$$

for some constant C where

$$T_s = \frac{\max_{|S_i \cap S_j| = s} \mathbb{E}\, \mathbb{I}_i \mathbb{I}_j}{n^s P(\mathcal{C})^2}. \tag{10}$$

To show that $Var(X)/(\mathbb{E}\, X)^2 \to 0$ it is enough to prove that $T_s \to 0$ for $s = 1, \ldots, h$.

Fix a positive integer s, $s \leq h$ and two sets $S_i, S_j \subseteq V$ of size h such that $|S_i \cap S_j| = s$. Let $\mathcal{C}^i = M(\mathcal{C}, S_i) = \{C_1^i, \ldots, C_t^i\}$ and $\mathcal{C}^j = M(\mathcal{C}, S_j) = \{C_1^j, \ldots, C_t^j\}$ be two copies of the diclique family $\mathcal{C}$ (here $t = |\mathcal{C}|$). Let M be a matching (not necessarily perfect) in a bipartite graph with parts $X = \{x_1, \ldots, x_t\}$ and $Y = \{y_1, \ldots, y_t\}$. Let $L = X \setminus \{x_l : x_l y_r \in M\}$ and $R = Y \setminus \{y_r : x_l y_r \in M\}$. Define a diclique family[1]

$$\mathcal{C}^M = \{C_l^i \cup C_r^j : x_l y_r \in M\} \cup \{C_l^i : x_l \in L\} \cup \{C_r^j : y_r \in R\}.$$

Here the union of diclique families is 'multiset union' so that $\mathcal{C}^M$ contains exactly $|M| + |R| + |L|$ elements.

Let us call M *good* if $\mathcal{C}^M[S^i] \setminus \mathcal{C}^i$ and $\mathcal{C}^M[S^j] \setminus \mathcal{C}^j$ consist of improper dicliques only.

Proposition 3.3 *If both $\mathcal{C}^i$ and $\mathcal{C}^j$ are induced in D then there is a good matching M such that D contains $\mathcal{C}^M$.*

Proof By definition, if both $\mathcal{C}^i$ and $\mathcal{C}^j$ are induced in D then there are two lists $w^i = (w_1^i, \ldots, w_t^i)$ and $w^j = (w_1^j, \ldots, w_t^j)$ of attributes from W such that

$$C(w_s^i)[S^i] = C_s^i \quad \text{and} \quad C(w_s^j)[S^j] = C_s^j \quad \text{for } s = 1, \ldots, t.$$

Also, there are no edges in $D[S^i]$ (respectively, $D[S^j]$) generated by attributes not in w^i (respectively, w^j). Notice also, that by definition the elements in each of the lists w^i, w^j must be distinct (although there can be some elements that belong to both lists). Thus all pairs $x_l y_r$ where $w_l^i = w_r^j$ and $l, r \in \{1, \ldots, t\}$ define some matching M_0. Clearly for each diclique $C \in \mathcal{C}_{M_0}$ it is possible to assign a unique element $w \in W$ such that $C \subseteq C(w)[S^i \cup S^j]$. The fact that M_0 is good follows by definition since both $\mathcal{C}^i$ and $\mathcal{C}^j$ are induced in D. $\square$

Let $P^{**}(\mathcal{C})$ denote the probability of the event that D contains $\mathcal{C}$. It follows by Proposition 3.3 that

$$\mathbb{E}\, \mathbb{I}_i \mathbb{I}_j \leq \sum_M P^{**}(\mathcal{C}^M). \tag{11}$$

[1] The notation should not be confused with the monochromatic cover $\mathcal{C}_M$.

where the sum is over all good matchings M. The number of all good matchings M is at most $(t+1)^t$, which is constant. By (10) and (11) it suffices to show that

$$\frac{P^{**}(\mathcal{C}^M)}{n^s P(\mathcal{C})^2} \to 0$$

for each good matching M. Write $S = S^i \cap S^j$. We have $P^{**}(\mathcal{C}^M) \le \tilde{P}(\mathcal{C}^M)$ by (6) and $P(\mathcal{C}_i) = P(\mathcal{C}_j) \sim \tilde{P}(\mathcal{C})$ by (2). For $x_l y_r \in M$ we have

$$\frac{\tilde{P}(C_l^i \cup C_r^j)}{\tilde{P}(C_l^i)\tilde{P}(C_r^j)} = \frac{1}{mp_-^{|C_l^{i-} \cap C_r^{j-}|} p_+^{|C_l^{i+} \cap C_r^{j+}|}} \le \frac{1}{mp_-^{|S \cap C_r^{j-}|} p_+^{|S \cap C_r^{j+}|}}.$$

Therefore we get that there is a constant c such that:

$$\begin{aligned}
\frac{P^{**}(\mathcal{C}^M)}{n^s P(\mathcal{C})^2} &= \frac{P^{**}(\mathcal{C}^M)}{n^s P(\mathcal{C}^i) P(\mathcal{C}^j)} \\
&\le \frac{c}{n^s} \prod_{x_l y_r \in M} \frac{\tilde{P}(C_l^i \cup C_r^j)}{\tilde{P}(C_l^i)\tilde{P}(C_r^j)} \prod_{x_l \in L} \frac{\tilde{P}(C_l^i)}{\tilde{P}(C_l^i)} \prod_{y_r \in R} \frac{\tilde{P}(C_r^j)}{\tilde{P}(C_r^j)} \\
&= \frac{c}{n^s} \prod_{x_l y_r \in M} \frac{1}{mp_-^{|C_l^{i-} \cap C_r^{j-}|} p_+^{|C_l^{i+} \cap C_r^{j+}|}} \\
&\le \frac{c}{n^s} \prod_{y_r \notin R} \frac{1}{mp_-^{|S \cap C_r^{j-}|} p_+^{|S \cap C_r^{j+}|}} \\
&= O\left(\frac{1}{\mathbb{E}\, X(\mathcal{C}^j[S])}\right) = o(1).
\end{aligned}$$

Here we get the bound in the last line as follows. If all dicliques in $\mathcal{C}^j[S]$ are proper, we use (3). If $\mathcal{C}^j[S]$ has some improper dicliques, by Lemma 3.1 it is sufficient to consider the case where the assumption (4) holds. In this case we use (5) with the family $\hat{\mathcal{C}}^j[S]$. This completes the proof. □

The following fact follows easily from the estimates above.

Lemma 3.4 *Fix an integer $h \ge 2$. Let H be a digraph with $V(H) = [h]$ and at least one arc. Let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that $mp_-p_+ \to 0$. Let $\mathcal{S}$ be a collection of simple diclique covers of H. Suppose that $\mu(\mathcal{C}) = O(1)$ for each $\mathcal{C} \in \mathcal{S}$ and for each simple cover $\mathcal{C}$ of H, such that $\mathcal{C} \notin \mathcal{S}$ we have $\mu(\mathcal{C}) \to 0$. Let $\mathcal{C}$ be any proper diclique cover of H such that $\mathcal{C} \notin \mathcal{S}$. Then $P(\mathcal{C} \in D(k)) \to 0$.*

Proof Suppose, the claim is false, i.e. there is a proper diclique cover $\mathcal{C}_0$ of H such that $\limsup P(\mathcal{C}_0 \in D(k)) > 0$ and $\mathcal{C}_0 \notin \mathcal{S}$. By assumption and

(7) $\mathcal{C}_0$ cannot be simple. We may assume that $\mathcal{C}_0$ consists of some simple cover $\mathcal{C}_1$ of H and a proper diclique C. By (7) and the definition of μ: $P(\mathcal{C}_0 \in D(k)) \le \mu(\mathcal{C}_0) \le \mu(\mathcal{C}_1)\tilde{P}(C) \to 0$ which is a contradiction. $\square$

4 Main proofs

The four special (see Section 2) diclique covers of the digraph $\overrightarrow{K}_h$ have the following birth thresholds:

Lemma 4.1 *Fix an integer $h \ge 2$. Let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that n is increasing, $p_- \to 0$, $p_+ \to 0$ and $mp_-p_+ \to 0$. Then the diclique covers $\mathcal{C}_M$, $\mathcal{C}_R$, $\mathcal{C}_{in}$, $\mathcal{C}_{out}$ of $\overrightarrow{K}_h$ have birth threshold functions $\tau_1, \tau_2, \tau_3, \tau_4$, respectively.*

Proof Let $\mathcal{C}$ be one of the four special covers. To prove that τ is a birth threshold function for $\mathcal{C}$ we may use Lemma 3.2. By that lemma it is enough to show that whenever $p_- \to 0$, $p_+ \to 0$ and $mp_-p_+ \to 0$ we have

$$\tau \to 0 \implies \mathbb{E}\, X(\mathcal{C}) \to 0; \tag{12}$$

$$\tau \to \infty \implies \text{for each non-empty set } S \subseteq [h] : \mathbb{E}\, X(\mathcal{C}[S]) \to \infty. \tag{13}$$

Here $\tau = \tau(k) = \tau(n(k), m(k), p_-(k), p_+(k))$.

By Lemma 3.1 we may assume that (4) holds.

For the monochromatic cover $\mathcal{C}_M$ we see that for any non-empty set $S \subseteq V(\mathcal{C}_M)$ of size $s \le h$, the restriction $\mathcal{C}_M[S]$ has the same form as the cover $\mathcal{C}_M$ defined for $h = s$ so:

$$\mu(\mathcal{C}_M[S]) = n^s m (p_-p_+)^s = m^{1-\frac{s}{h}} \tau_1^s.$$

So (12) and (13) follow by (3).

Consider now the diclique cover $\mathcal{C}_R$. Let $S \subseteq V(\mathcal{C}_R)$ be non-empty and write $s = |S|$. When $s < h$ the restriction $\mathcal{C} = \mathcal{C}_R[S]$ is not a proper diclique family as for each $v \in S$ it contains $(h-s)$ dicliques $(\{v\};\emptyset)$ and $(h-s)$ dicliques $(\emptyset;\{v\})$. Therefore we need to use (5) with the family $\hat{\mathcal{C}}$.

If $mp_- \to 0$ and $mp_+ \to 0$ or $S = [h]$ then $\hat{\mathcal{C}} = \mathcal{C}$ and

$$\tilde{\mu}(\mathcal{C}) = n^s m^{s(s-1)+2s(h-s)} (p_-p_+)^{s(s-1)+s(h-s)} = m^{s(h-s)} \tau_2^{s(h-1)}$$

and (12) follows by (3). If $mp_- \to a_1 > 0$ and $mp_+ \to 0$ then $\hat{\mathcal{C}} = \mathcal{C} \setminus \{(\{v\};\emptyset) : v \in S\}$.

$$\begin{aligned}\tilde{\mu}(\mathcal{C}) &= n^s m^{s(s-1)+2s(h-s)} (p_-p_+)^{s(s-1)+s(h-s)} (mp_-)^{-s(h-s)} \\ &= (p_-)^{-s(h-s)} \tau_2^{s(h-1)}.\end{aligned}$$

The expression $\tilde{\mu}(\mathcal{C})$ for the case $mp_+ \to b_1 > 0$ and $mp_- \to 0$ is similar (replace p_- with p_+).

If $mp_- \to a_1 > 0$ and $mp_+ \to b_1 > 0$ by (5) the family $\hat{\mathcal{C}}$ is exactly the cover $\mathcal{C}_R$ defined for $h = s$ and

$$\tilde{\mu}(\mathcal{C}) = n^{\frac{s(h-s)}{h-1}} \tau_2^{s(s-1)}.$$

In each case (13) holds by (5).

For the "in-stars" cover $\mathcal{C}_{in}$ we have

$$\mu(\mathcal{C}_{in}) = n^h m^h p_-^{h(h-1)} p_+^h = \tau_3^h$$

so the implication (12) holds by (3). Now let $S \subset V(\mathcal{C}_{in})$ be non-nempty and write $s = |S|$. Suppose $s \geq 2$. Then the diclique family $\mathcal{C} = \mathcal{C}_{in}[S]$ is not proper as it contains $h - s$ copies of the diclique $(S; \emptyset)$. If $mp_-^s \to 0$, we use (5) with $\hat{\mathcal{C}} = \mathcal{C}$:

$$\tilde{\mu}(\mathcal{C}) = n^s m^{s+(h-s)} p_-^{s(s-1)+s(h-s)} p_+^s = m^{h-s} \tau_3^s.$$

If $mp_-^s \to a_s > 0$ then we have $\hat{\mathcal{C}} = \{(S \setminus \{v\}; \{v\}) : v \in S\}$ so

$$\tilde{\mu}(\mathcal{C}) = n^s m^{s+(h-s)} p_-^{s(s-1)+s(h-s)} p_+^s (mp_-^s)^{-(h-s)} = p_-^{-s(h-s)} \tau_3^s.$$

Now, if $s = 1$, in each of the cases $a_1 > 0$, $b_1 > 0$ and $a_1 = b_1 = 0$, see (4), by (5) we have that $\mathbb{E}\, X(\mathcal{C}) = \Omega(\tau_3)$. Therefore by (5) we see that

$$\mathbb{E}\, X(\mathcal{C}) \to \infty \tag{14}$$

when $\tau_3 \to \infty$ and $S \subseteq V(\mathcal{C}_{in})$ is non-empty and so (13) holds for $\mathcal{C}_{in}$. Finally, the case of $\mathcal{C}_{out}$ is symmetric to that of $\mathcal{C}_{in}$. $\square$

Remark 4.2 *Let $\{D(k)\}$ be as in Lemma 4.1. Let $\mathcal{C}$ be one of the four special diclique covers of $\overrightarrow{K}_h$ and let τ be its birth threshold function given in Lemma 4.1. If $\tau = \Theta(1)$ then $P($a copy of $\mathcal{C}$ is induced in $D(k)) = \Omega(1)$.*

Proof From the proof of Lemma 4.1 we have that $\mu(\mathcal{C}) = \Theta(\tau^h)$ for $\mathcal{C} \in \{\mathcal{C}_{in}, \mathcal{C}_M, \mathcal{C}_{out}\}$ and $\mu(\mathcal{C}_R) = \Theta(\tau_2^{h(h-1)})$. So in each case, we have $\mu(\mathcal{C}) = \Theta(1)$.

By (3) the number $X = X(\mathcal{C})$ of induced copies of $\mathcal{C}$ satisfies

$$\mathbb{E}\, X = \Theta(\mu(\mathcal{C})) = \Theta(1).$$

The proof of Lemma 4.1 also shows that in each case $\mu(\mathcal{C}[S]) = \Omega(1)$ for each $S \subseteq V(C)$. Following the lines of the second part of the proof of Lemma 3.2 we see that

$$\frac{Var(X)}{(\mathbb{E}\, X)^2} = O(1).$$

Using the Cauchy-Schwartz inequality we get

$$P(X > 0) \geq \frac{(\mathbb{E}\, X)^2}{\mathbb{E}\, X^2} = \Omega(1).$$

□

The next lemma says that if p_- is sufficiently large then it is always 'better' to replace any diclique cover by a 'star' cover $\mathcal{C}'$:

Lemma 4.3 *Let $h \geq 3$ be an integer and let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that n is increasing, $p_- \to 0$ and $m^{\frac{h-1}{h(h-2)}} p_- = \Omega(1)$. Let $\mathcal{C}$ be a diclique family consisting of a single proper diclique $C = (C^-; C^+)$ with $V(C) \subseteq [h]$. Suppose $|C^+| \geq 2$ or $|C^- \cap C^+| = 1$. Let $\mathcal{C}'$ be a diclique family defined by*

$$\mathcal{C}' = \left\{ (C^- \setminus \{v\}; \{v\}) \,:\, v \in C^+ \right\}.$$

Then $\tilde{P}(\mathcal{C}') = \Omega(\tilde{P}(\mathcal{C}))$. More precisely,

(a) If $\mathcal{C} \neq \mathcal{C}_M$ or $m^{\frac{h-1}{h(h-2)}} p_- \to \infty$ then $\tilde{P}(\mathcal{C}) = o(\tilde{P}(\mathcal{C}'))$;

(b) If $\mathcal{C} = \mathcal{C}_M$ and $m^{\frac{h-1}{h(h-2)}} p_- = \Theta(1)$ then $\tilde{P}(\mathcal{C}') = \Theta(\tilde{P}(\mathcal{C}))$.

We note that $\mathcal{C}$ and $\mathcal{C}'$ are both diclique covers of the diclique C.
Proof If $|C^+| = 1$ and $|C^- \cap C^+| = 1$ then

$$\frac{\tilde{P}(\mathcal{C}')}{\tilde{P}(\mathcal{C})} = p_-^{-1} \to \infty$$

so we may assume that $|C^+| > 1$. Let $c_- = |C^-|$ and $c_+ = |C^+|$. Then

$$\begin{aligned} \frac{\tilde{P}(\mathcal{C}')}{\tilde{P}(\mathcal{C})} &= \frac{m^{c_+} p_-^{\sum \mathcal{C}'^-} p_+^{c_+}}{m p_-^{c_-} p_+^{c_+}} = m^{c_+ - 1} p_-^{\sum \mathcal{C}'^- - c_-} \\ &\geq m^{c_+ - 1} p_-^{(c_+ - 1)(h-1) - 1} = \left(m p_-^{h - 1 - \frac{1}{c_+ - 1}} \right)^{c_+ - 1} \qquad (15) \\ &\geq \left(m p_-^{h - 1 - \frac{1}{h-1}} \right)^{c_+ - 1}. \qquad (16) \end{aligned}$$

Here the inequality in the second line follows from:

$$\sum \mathcal{C}'^{-} - c_{-} \leq (c_{+} - 1)(h-1) - 1. \tag{17}$$

To see (17) consider three possible cases:

If $c_{-} \in \{1, 2, \ldots, h-2\}$ then

$$\sum \mathcal{C}'^{-} - c_{-} \leq c_{-}c_{+} - c_{-} = c_{-}(c_{+} - 1) \leq (h-1)(c_{+}-1) - (c_{+}-1).$$

If $c_{-} = h-1$ then there is at most one diclique C' in $\mathcal{C}'$ with $C'^{+} = \{v\}$ such that $v \in |C^{+} \setminus C^{-}|$ so

$$\sum \mathcal{C}'^{-} - c_{-} \leq (c_{-}-1)(c_{+}-1) + c_{-} - c_{-} = (h-1)(c_{+}-1) - (c_{+}-1),$$

and if $c_{-} = h$,

$$\sum \mathcal{C}'^{-} - c_{-} \leq c_{+}(c_{-}-1) - c_{-} = (h-1)(c_{+}-1) - 1.$$

Note that in the inequality (16) the right hand side can be of the same order only for $c_{+} = h$. But if $c_{+} = h$ and $c_{-} < h$, we get

$$\sum \mathcal{C}'^{-} - c_{-} = (h-2)c_{-} < h(h-2) = c_{+}(h-1) - 1,$$

so in that case the right hand side of the inequality (15) of a smaller order than the left hand side. Now note that

$$mp_{-}^{h-1-\frac{1}{h-1}} = \Omega(1)$$

since $p_{-} = \Omega\left(m^{-\frac{h-1}{h(h-2)}}\right)$ and $m^{1-\frac{h-1}{h(h-2)}\left(h-1-\frac{1}{h-1}\right)} = 1$. Thus we have shown that $\tilde{P}(\mathcal{C}) = O(\tilde{P}(\mathcal{C}'))$ and the claim (a) holds. To complete the proof, note that for $\mathcal{C} = \mathcal{C}_M$ (17) and (16) become equalities. □

The next lemma shows that if we have a diclique family consisting of many "in-stars" centred at one vertex, we may merge all of them into a single diclique.

Lemma 4.4 *Let $h \geq 2$ be an integer and let H be a digraph obtained from the complete bipartite graph $K_{1,h-1}$ by orienting each edge towards the centre vertex v. Let $\mathcal{C}$ be any proper diclique cover of the digraph H of the form*

$$\mathcal{C} = \left\{(C_i^{-}; \{v\}),\, i = 1, \ldots, t\right\}$$

where $t \geq 2$ and $\bigcup_i C_i^- = V(H) \setminus \{v\}$.

Let $\mathcal{C}^* = \{(V(H) \setminus \{v\}; \{v\})\}$ be a cover of H and let $\{D(k)\}$ be a sequence of random binomial intersection digraphs such that n is increasing, $p_- \to 0$ and $mp_+ = O(1)$. Then

$$\tilde{P}(\mathcal{C}^*) = \Omega(\tilde{P}(\mathcal{C})).$$

More precisely,

(a) If $mp_+ = o(1)$ or $\sum \mathcal{C}^- > h - 1$ then $\tilde{P}(\mathcal{C}) = o(\tilde{P}(\mathcal{C}^*))$;

(b) If $mp_+ = \Theta(1)$ and $\sum \mathcal{C}^- = h - 1$ then $\tilde{P}(\mathcal{C}^*) = \Theta(\tilde{P}(\mathcal{C}))$.

Proof Using (2) we have

$$\frac{\tilde{P}(\mathcal{C})}{\tilde{P}(\mathcal{C}^*)} = \frac{a_{\mathcal{C}} m^t p_-^{\sum \mathcal{C}^-} p_+^t}{m p_-^{h-1} p_+} = a_{\mathcal{C}} (mp_+)^{t-1} p_-^{\sum \mathcal{C}^- - (h-1)} = O(1)$$

since $\sum \mathcal{C}^- \geq |\bigcup_i C_i| = h - 1$. The claims (a) and (b) follow similarly. $\square$

Let us remark that we use $\tilde{P}$ rather than P in Lemma 4.3 and Lemma 4.4 for convenience. By (2) we know that P can be replaced by $\tilde{P}$ as long as p_-, p_+ and mp_-p_+ all tend to 0. We are now ready to prove Theorem 2.1.

We split the proof of our main result into a few lemmas. In the four lemmas below we assume that $h, \{D(k)\}, n, \alpha_0, m$ are as in Theorem 2.1 $p_- \to 0, p_+ \to 0$ and $\mu(\mathcal{C}) = \mu(\mathcal{C}, n, m, p_-, p_+)$ is as defined in Section 3.

Lemma 4.5 *Suppose $\alpha < \alpha_0$ and $m^{\frac{h-1}{h(h-2)}} p_- = \Omega(1)$. Then τ_3 is a birth threshold function for $\overrightarrow{K}_h$. Furthermore, if $m^{\frac{h-1}{h(h-2)}} p_- \to \infty$ then the leading set is $\mathcal{L} = \{\mathcal{C}_{in}\}$ and if $m^{\frac{h-1}{h(h-2)}} p_- = \Theta(1)$ then the leading set is $\mathcal{L} = \{\mathcal{C}_{in}, \mathcal{C}_M\}$.*

In each of the cases above, if $\tau_3 = \Theta(1)$ then $\mu(\mathcal{C}) \to 0$ for any simple diclique cover $\mathcal{C} \notin \mathcal{L}$ and $\mu(\mathcal{C}) = \Theta(1)$ for $\mathcal{C} \in \mathcal{L}$.

Lemma 4.6 *Suppose $\alpha < \alpha_0$, $m^{\frac{h-1}{h(h-2)}} p_- \to 0$ and $m^{\frac{h-1}{h(h-2)}} p_+ \to 0$. Then τ_1 is a birth threshold function for $\overrightarrow{K}_h$ with the leading set $\mathcal{L} = \{\mathcal{C}_M\}$.*

If $\tau_1 = \Theta(1)$ then $\mu(\mathcal{C}) \to 0$ for any simple diclique cover $\mathcal{C} \notin \mathcal{L}$ and $\mu(\mathcal{C}) = \Theta(1)$ for $\mathcal{C} \in \mathcal{L}$.

Lemma 4.7 *Suppose $\alpha \geq \alpha_0$ and $mp_+ = O(1)$. Then τ_3 is a birth threshold function for $\overrightarrow{K}_h$. Furthermore, if $\alpha \geq \alpha_0$ and $mp_+ \to 0$ then the leading set is $\mathcal{L} = \{\mathcal{C}_{in}\}$; if $\alpha > \alpha_0$ and $mp_+ = \Theta(1)$ then the leading set is $\mathcal{L} = \mathcal{S}_{in}$; if $\alpha = \alpha_0$ and $mp_+ = \Theta(1)$ then the leading set is $\mathcal{L} = \{\mathcal{C}_M\} \cup \mathcal{S}_{in}$.*

In each of the cases above, if $\tau_3 = \Theta(1)$ then $\mu(\mathcal{C}) \to 0$ for any simple diclique cover $\mathcal{C} \notin \mathcal{L}$ and $\mu(\mathcal{C}) = \Theta(1)$ for $\mathcal{C} \in \mathcal{L}$.

Lemma 4.8 *Suppose $\alpha \geq \alpha_0$, $mp_- \to \infty$ and $mp_+ \to \infty$. Then τ_2 is a birth threshold function for $\overrightarrow{K}_h$. Furthermore, if $\alpha > \alpha_0$ then the leading set is $\mathcal{L} = \{\mathcal{C}_R\}$ and if $\alpha = \alpha_0$ then the leading set is $\mathcal{L} = \{\mathcal{C}_M, \mathcal{C}_R\}$.*

In each of the cases above, if $\tau_2 = \Theta(1)$ then $\mu(\mathcal{C}) \to 0$ for any simple diclique cover $\mathcal{C} \notin \mathcal{L}$ and $\mu(\mathcal{C}) = \Theta(1)$ for $\mathcal{C} \in \mathcal{L}$.

Proof of Theorem 2.1 Apply Lemmas 4.5-4.8 and notice that the cases (i)(c) and (ii)(c) follow by symmetry. □

Proof of Theorem 2.2 We note that if τ, τ' are birth threshold functions for the sequence $\{D(k)\}$ given by Theorem 2.2 then they are equivalent in the sense that $\log \tau = \Theta(\log \tau')$. The cases (a), (c), (d), (f) follow by Lemmas 4.5-4.8 and the remaining cases follow by symmetry. □

The idea of the proof of Lemmas 4.5-4.8 is to consider the birth threshold functions of the four special diclique covers and the boundaries determined by them, see Figure 2. We will use Lemma 4.3 and Lemma 4.4 to compare the probability of complicated diclique covers of $\overrightarrow{K}_h$ with the probability of appropriate special covers. In the proofs we write $D = D(k)$.

Proof of Lemma 4.5 By Lemma 3.1 we may assume that $mp_-p_+ \to a \in [0; \infty]$.

Suppose that $\tau_3 = \tau_3(k) = nmp_-^{h-1}p_+ \to \infty$ as $k \to \infty$. If $a = 0$ then by Lemma 4.1 the random digraph D contains a copy of $\mathcal{C}_{in}$ whp. If $a > 0$ consider a sequence of random digraphs $\{D'(k)\}$ where $D'(k) = D(n, m, p_-, p'_+)$, $p'_+ = (\omega mp_-)^{-1}$ and $\omega = \omega(n)$ grows slowly, say $\omega(n) = \ln n$. We have

$$\tau_3(n, m, p_-, p'_+) = nmp_-^{h-1}p'_+ = np_-^{h-2}\omega^{-1} = \Omega\left(n^{1-\frac{\alpha(h-1)}{h}}\omega^{-1}\right).$$

So $\tau_3(n, m, p_-, p'_+) \to \infty$, $mp_-p'_+ \to 0$ and $p'_+ = o(p_+)$. We have that $D'(k)$ contains a copy of $\mathcal{C}_{in}$ whp by Lemma 4.1 and therefore $D(k)$ contains a copy of $\mathcal{C}_{in}$ whp by monotonicity.

Now suppose that $\tau_3 = \Theta(1)$, so that $\mu(\mathcal{C}_{in}) = \tau_3^h = \Theta(1)$. Notice that since $\alpha < 1$ we have

$$mp_-p_+ = \frac{\tau_3}{np_-^{h-2}} = O\left(\frac{1}{nm^{-\frac{h-1}{h}}}\right) = O\left(n^{-1+\frac{\alpha(h-1)}{h}}\right) = o(1).$$

Let $\mathcal{C}$ be any simple diclique cover of $\overrightarrow{K}_h$. We will show that

$$\mu(\mathcal{C}) = O(1) \tag{18}$$

and furthermore we can replace $O()$ with $o()$ (respectively, $\Theta()$) if $\mathcal{C} \notin \mathcal{L}$ (respectively, $\mathcal{C} \in \mathcal{L}$).

Assuming we have proved (18), we obtain that $\tau_3 \to 0$ implies $\mu(\mathcal{C}) \to 0$, since both τ_3 and $\mu(\mathcal{C})$ are increasing multinomials in m, p_-, p_+. Now the fact that τ_3 is a birth threshold function for $\overrightarrow{K}_h$ follows by (6) and Markov's inequality since the number of simple diclique covers of $\overrightarrow{K}_h$ is finite. Remark 4.2 implies that in the case $\mu(\mathcal{C}_{in}) = \tau_3^h = \Theta(1)$ we have $P(\mathcal{C}_{in} \in D) = \Omega(1)$. If in addition $mp_-^{\frac{h-1}{h(h-2)}} = \Theta(1)$ then we have $\mu(\mathcal{C}_M) = \tau_1(n, m, p_-, p_+)^h = \Theta(1)$ and so $P(\mathcal{C}_M \in D) > 0$. So by Lemma 3.4 the set $\mathcal{L}$ is leading (for the birth threshold function τ_3 and $\overrightarrow{K}_h$).

So let us prove (18). Suppose $\mathcal{C} = \{C_1, C_2, \ldots, C_t\}$. Using the definition of μ and $\tilde{P}$, see Section 3, we have

$$\begin{aligned}
\mu(\mathcal{C}) &= (h!)^{-1} n^h \tilde{P}(C_1)\tilde{P}(C_2)\ldots\tilde{P}(C_t) \\
&= O\left(n^h \tilde{P}(\mathcal{C}_1')\tilde{P}(\mathcal{C}_2')\ldots\tilde{P}(\mathcal{C}_t')\right) && (19)\\
&= O(n^h \tilde{P}(\mathcal{C}_{in})) && (20)\\
&= O(\mu(\mathcal{C}_{in})) = O(1). && (21)
\end{aligned}$$

Here in (19) we apply Lemma 4.3 so that for each $i = 1, \ldots, t$ the diclique family $\mathcal{C}_i'$ is obtained from the family $\mathcal{C}_i = \{C_i\}$ by splitting it into "in-stars". The resulting family $\mathcal{C}' = \mathcal{C}_1' \cup \mathcal{C}_2' \cup \cdots \cup \mathcal{C}_t'$ is a cover of $\overrightarrow{K}_h$ which consists of possibly duplicated or overlapping "in-stars". Next, in (20) we regroup the terms and apply Lemma 4.4.

More precisely, note that by Lemma 4.3 we can replace $O()$ with $o()$ in (19) in all cases except if

- $p_-m^{\frac{h-1}{h(h-2)}} = \Theta(1)$ and $\mathcal{C} = \mathcal{C}_M$; or
- for each $C \in \mathcal{C}$ we have $|C^+| = 1$ and $C^+ \cap C^- = \emptyset$.

We can complete the proof of (18) by noticing that since $\alpha < \alpha_0$ we have $p_+ = \Theta\left(\tau_3/(mnp_-^{h-1})\right) = o(m^{-1})$, so by Lemma 4.4 we can replace $O()$ with $o()$ in (20) for the second exceptional case whenever $\mathcal{C} \neq \mathcal{C}_{in}$. $\square$

Proof of Lemma 4.6 First suppose that $\tau_1 = \tau_1(k) = nm^{1/h}p_-p_+ \to \infty$. By Lemma 4.1, some copy of $\mathcal{C}_M$ is induced in D whenever $mp_-p_+ \to 0$. Otherwise, using Lemma 3.1 we may assume that $mp_-p_+ \to a \in (0;\infty]$. Let $p'_- = (mp_+ \ln n)^{-1}$ and consider a sequence of random digraphs $\{D'(k)\}$ where $D'(k) = D(n, m, p'_-, p_+)$. Then $mp'_-p_+ \to 0$, $p'_- = o(p_-)$ and $\tau_1(n, m, p'_-, p_+) \to \infty$. So we have that $D'(k)$ contains an induced copy of $\mathcal{C}_M$ whp by Lemma 4.1 and by monotonicity $\mathcal{C}_M \in D$ whp.

Now suppose that $\tau_1 = \Theta(1)$. In this case we have $\mu(\mathcal{C}_M) = \tau_1^h = \Theta(1)$ and

$$mp_-p_+ = O\left(m\frac{1}{nm^{1/h}}\right) = O\left(\frac{m^{\frac{h-1}{h}}}{n}\right) = O\left(n^{-1+\alpha\frac{h-1}{h}}\right) \to 0.$$

Let us now show that for any simple diclique cover $\mathcal{C}$ of $\overrightarrow{K}_h$, such that $\mathcal{C} \neq \mathcal{C}_M$ we have $\mu(\mathcal{C}) \to 0$. Lemma 4.6 will then follow by monotonicity, Lemma 3.4, Lemma 4.1 and Remark 4.2.

Write $t = p_-p_+ = \Theta\left(n^{-1}m^{-1/h}\right)$. By the definition of μ, see (3), we have that for any simple diclique cover $\mathcal{C}$ of $\overrightarrow{K}_h$

$$\begin{aligned} h!\mu(\mathcal{C}) &= n^h m^{|\mathcal{C}|} p_+^{\sum \mathcal{C}^+} p_-^{\sum \mathcal{C}^-} \\ &= n^h m^{|\mathcal{C}|} p_+^{\sum \mathcal{C}^+ - \sum \mathcal{C}^-} t^{\sum \mathcal{C}^-} && (22) \\ &= n^h m^{|\mathcal{C}|} p_-^{\sum \mathcal{C}^- - \sum \mathcal{C}^+} t^{\sum \mathcal{C}^+}. && (23) \end{aligned}$$

Define $p_0 = m^{-\frac{h-1}{h(h-2)}}$. Clearly, if $\sum \mathcal{C}^- \geq \sum \mathcal{C}^+$, then by (23):

$$\mu(\mathcal{C}) = \mu(\mathcal{C}, n, m, p_-, p_+) = O\left(\mu(\mathcal{C}, n, m, p_0; t/p_0)\right).$$

Since $\tau_3(n, m, p_0, t/p_0) = O(1)$ by Lemma 4.6 we get $\mu' = \mu(\mathcal{C}, n, m, p_0, t/p_0) = O(1)$ and furthermore $\mu' \to 0$ if $\mathcal{C} \notin \{\mathcal{C}_M, \mathcal{C}_{in}\}$.

To see why $\mu(\mathcal{C}_{in}) \to 0$, note that $\sum \mathcal{C}_{in}^- > \sum \mathcal{C}_{in}^+$ and by (23) we have $\mu(\mathcal{C}_{in}, n, m, p_0; t/p_0) = o\left(\mu(\mathcal{C}_{in}, n, m, p_-, p_+)\right)$.

The case $\sum \mathcal{C}^+ > \sum \mathcal{C}^-$ is similar by symmetry: we use Lemma 4.5 with p_- and p_+ interchanged and replace τ_3 with τ_4, $\mathcal{C}_{in}$ with $\mathcal{C}_{out}$. $\square$

Proof of Lemma 4.7 Suppose $\tau_3 \to \infty$. If $mp_-p_+ \to 0$ then by Lemma 4.1, D contains a copy of $\mathcal{C}_{in}$ whp. Otherwise let $\omega \to \infty$ not too fast so that $\omega =$

$o\left(\tau_3^{1/(h-1)}\right)$ and let $p_-' = p_-/\omega$. Consider a sequence of random digraphs $\{D'(k)\}$ where $D'(k) = D(n, m, p_-', p_+)$. By Lemma 4.1 and Lemma 3.1 $P(\mathcal{C}_{in} \in D'(k)) \to 1$ and so $P(\mathcal{C}_{in} \in D) \to 1$ by monotonicity.

Now suppose $\tau_3 = \Theta(1)$, so that $\mu(\mathcal{C}_{in}) = \tau_3^h = \Theta(1)$. In this case we have $mp_-p_+ = O(n^{-1/(h-1)}) \to 0$. Let $\mathcal{C}$ be a any simple diclique cover of $\overrightarrow{K}_h$. We will show that

$$\mu(\mathcal{C}) = O(1) \tag{24}$$

and furthermore we can replace $O()$ with $o()$ (respectively, $\Theta()$) if $\mathcal{C} \notin \mathcal{L}$ (respectively, $\mathcal{C} \in \mathcal{L}$).

Assuming (24) holds, since both τ_3 and $\mu(\mathcal{C})$ are monotone increasing multinomials in m, p_-, p_+ we have that $\tau_3 \to 0$ implies $\mu(\mathcal{C}) \to 0$. Therefore using Lemma 3.4, Lemma 4.1 and Remark 4.2 completes the proof of the lemma in the case where $mp_- \to 0$. For the boundary case $mp_- = \Theta(1)$, (24) and Lemma 4.9 below shows that each cover in $\mathcal{S}_{in}$ belongs to the leading set when $\alpha \geq \alpha_0$. If $\alpha = \alpha_0$ we use the fact that $\mu_1(\mathcal{C}_M) = \tau_1^h = \Theta(1)$ and Remark 4.2 to show that $\mathcal{C}_M$ also belongs to $\mathcal{L}$.

Let us check (24). We have

$$p_- = \left(\frac{\tau_3}{nmp_+}\right)^{1/(h-1)} = \Omega\left(m^{\frac{-1}{\alpha(h-1)}}\right) = \Omega\left(m^{-\frac{h-1}{h(h-2)}}\right)$$

and $m^{\frac{h-1}{h(h-2)}} p_- \to \infty$ when $\alpha > \alpha_0$. So we may apply Lemmas 4.3 and 4.4 as in the proof of Lemma 4.5 to get that

$$\mu(\mathcal{C}) = O(\mu(\mathcal{C}_{in})) = O(1) \tag{25}$$

Furthermore, Lemma 4.3 and Lemma 4.4 also give that we may replace $O()$ with $o()$ in (25) in all cases, except if

- $\mathcal{C} = \mathcal{C}_M$ and $m^{\frac{h-1}{h(h-2)}} p_- = \Theta(1)$ or
- $mp_+ = \Theta(1), |C^+| = 1$ for each $C \in \mathcal{C}$ and for each $j = 1, \ldots, h$ the diclique cover $\mathcal{C}_j$ obtained by taking all dicliques $C \in \mathcal{C}$ that have $C^+ = \{j\}$ satisfies $\sum \mathcal{C}_j^- = h - 1$.

The first exception occurs only if $\alpha = \alpha_0$ and $mp_+ = \Theta(1)$. The second exception represents all diclique covers $\mathcal{C} \in \mathcal{S}_{in}$. This completes the proof of (24). $\square$

Proof of Lemma 4.8 Suppose $\tau_2 = n^{1/(h-1)} mp_-p_+ \to \infty$. If $mp_-p_+ \to 0$ the random digraph D contains a copy of $\mathcal{C}_R$ whp by Lemma 4.1. Otherwise, assume that $mp_-p_+ \to a \in (0; \infty]$. Let $p_-' = (mp_+ \ln n)^{-1} = o(p_-)$.

Then the random digraph $D(n,m,p'_-,p_+)$ contains a copy of $\mathcal{C}_R$ whp by Lemma 4.1 since $mp'_-p_+ \to 0$ and $\tau_2 \to \infty$. Monotonicity and Lemma 3.1 imply that D contains a copy of $\mathcal{C}_R$ whp in all cases when $\tau_2 \to \infty$.

Now suppose that $\tau_2 = \Theta(1)$. Similarly as above, we have $\mu(\mathcal{C}_R) = \tau_2^{h-1} = \Theta(1)$ and $mp_-p_+ = O\left(n^{-1/(h-1)}\right) \to 0$. Let $\mathcal{C}$ be a simple diclique cover of $\overrightarrow{K}_h$. We will show that if $\mathcal{C} \not\cong \mathcal{L}$ then $\mu(\mathcal{C}) \to 0$. As in the proof of Lemma 4.6 we will then be able to complete the proof using Lemma 4.1, Markov's inequality, Lemma 3.4 and Remark 4.2 (for the case $\alpha = \alpha_0$ notice that $\tau_1(n,m,p_-,p_+) = \Theta(1)$, so the birth threshold functions τ_1 and τ_2 become equivalent).

Suppose $\sum \mathcal{C}^- \geq \sum \mathcal{C}^+$. Using (22) we have

$$\mu(\mathcal{C}) = \mu(\mathcal{C},n,m,p_-,p_+) = O\left(\mu(\mathcal{C},n,m,t/p_0,p_0)\right)$$

where $t = p_-p_+$ and $p_0 = m^{-1}$. Also note that when $\mathcal{C}^- > \mathcal{C}^+$

$$\mu(\mathcal{C}) = o(\mu(\mathcal{C},n,m,t/p_0,p_0)). \tag{26}$$

Since $\tau_3(n,m,t/p_0,p_0) = nm\left(n^{-1/(h-1)}m^{-1}p_0^{-1}\right)^{h-1}p_0 = \Theta(1)$ we can apply (25) from Lemma 4.7 for the sequence of random digraphs $\{D(n,m,t/p_0,p_0)\}$ to get that $\mu(\mathcal{C},n,m,t/p_0,p_0) = O(1)$.

Lemma 4.7 also gives that $\mu(\mathcal{C},n,m,t/p_0,p_0) \to 0$ for all simple diclique covers, except if $\mathcal{C} = \mathcal{C}_M$ in the case $\alpha = \alpha_0$ or if $\mathcal{C} \in \mathcal{S}_{in}$. It remains to check that $\mu(\mathcal{C}_0) \to 0$ for any diclique cover $\mathcal{C}_0 \in \mathcal{S}_{in}$ such that $\mathcal{C}_0 \neq \mathcal{C}_R$. But any $\mathcal{C}_0 \in \mathcal{S}_{in} \setminus \{\mathcal{C}_R\}$ has $\sum \mathcal{C}_0^- > \sum \mathcal{C}_0^+$, therefore $\mu(\mathcal{C}_0) \to 0$ by (26).

The case $\sum \mathcal{C}^- < \sum \mathcal{C}^+$ is similar because of symmetry. $\square$

Lemma 4.9 *Let $h, \{D(k)\}, n, \alpha_0, m$ be as in Theorem 2.1. Suppose $p_- \to 0$ and $p_+ \to 0$, $\alpha \geq \alpha_0$ and $mp_+ = \Theta(1)$. Consider any diclique cover $\mathcal{C}_0 \in \mathcal{S}_{in}$. If $\tau_3 = \tau_3(n,m,p_-,p_+) = \Omega(1)$ then $P(\mathcal{C} \in D(k)) = \Omega(1)$ and if $\tau \to \infty$ then $D(k)$ contains $\mathcal{C}$ whp.*

Proof Write $D = D(k)$. By Lemma 4.7 we have $\mu(\mathcal{C}_0) = \Theta(1)$ whenever $\tau_3 = \Theta(1)$.

Assume first that $\tau_3 \to \infty$. We claim that for any non-empty set $S \subseteq [h]$:

$$\mu(\mathcal{C}_0[S]) = \Omega(\mu(\mathcal{C}_{in}[S])). \tag{27}$$

By (14) in the proof of Lemma 4.1 we have that $\mu(\mathcal{C}_{in}[S]) \to \infty$.

By Lemma 3.1 we may assume that (4) holds. Also, since $p_- \to 0$ we have $mp_-p_+ \to 0$.

First consider the case $S = \{v\}$ for some $v \in [h]$. By the definition of $\mathcal{S}_{in}$, $\mathcal{C}_0[S]$ consists of exactly $h-1$ dicliques $(\{v\}, \emptyset)$ and one or more dicliques $(\emptyset; \{v\})$. Let us apply (5). Since $mp_+ = \Theta(1)$, the set $\hat{\mathcal{C}}_0[S]$ is equal to $\hat{\mathcal{C}}_{in}[S]$, so (27) follows by (14) for the case $|S| = 1$.

Now suppose $|S| \geq 2$. Split a given diclique family $\mathcal{C}$ into the family of its proper dicliques $\mathcal{C}'$ and the family of its improper dicliques $\mathcal{C}''$ so that $\mathcal{C} = \mathcal{C}' \cup \mathcal{C}''$. Note that if $V(\mathcal{C}'') \subseteq V(\mathcal{C}')$ by (5) we have

$$\mathbb{E}\, X(\mathcal{C}) = \Theta(\mu(\mathcal{C}')\tilde{P}(\tilde{\mathcal{C}}'')) \tag{28}$$

where $\tilde{\mathcal{C}}''$ is the collection of dicliques C of $\mathcal{C}''$ that satisfy $\tilde{P}(C) \to 0$.

By Lemma 4.4 we have $\mu(\mathcal{C}_0[S]') = \Theta(\mu(\mathcal{C}_{in}[S]'))$. If $mp_-^s \to 0$ then $\tilde{\mathcal{C}}_{in}[S]''$ consists of $h-s$ dicliques $(S; \emptyset)$. Similarly, since $\mathcal{C}_0 \in \mathcal{S}_{in}$, the set $\tilde{\mathcal{C}}_0[S]''$ can be partitioned into $h-s$ families of improper dicliques $\mathcal{C}_1, \ldots, \mathcal{C}_{h-s}$ so that each $\mathcal{C}_i$ consists of pairwise disjoint dicliques, $\sum \mathcal{C}_i^- \leq s$ and $\sum \mathcal{C}_i^+ = 0$. We have for each $i = 1, \ldots, h-s$:

$$\prod_{C \in \mathcal{C}_i} \tilde{P}(C) = m^{|\mathcal{C}_i|} p_-^{\sum \mathcal{C}_i^-} = \Omega(\tilde{P}\,(S; \emptyset)).$$

If $mp_-^s \to a_s > 0$ then both $\tilde{\mathcal{C}}_0[S]''$ and $\tilde{\mathcal{C}}_{in}[S]''$ are empty. So in each case

$$\tilde{P}(\tilde{\mathcal{C}}_0[S]'') = \Theta\left(\prod_i \tilde{P}(\mathcal{C}_i)\right) = \Omega(\tilde{P}(\tilde{\mathcal{C}}_{in}[S]''))$$

and (27) follows by (28). Now (27) and Lemma 3.2 imply that D contains a copy of $\mathcal{C}_0$ whp when $\tau_3 \to \infty$.

Finally, we need to check that $P(\mathcal{C}_0 \in D) = \Omega(1)$ if $\tau_3 = \Theta(1)$. We note that in this case (27) still holds. Therefore the same argument as in the proof of Remark 4.2 shows that $P(\mathcal{C}_0 \in D) = \Omega(1)$. □

Proof of Remark 2.3 Write $D = D(k)$. Consider two simple diclique covers of $\overrightarrow{K}_2$, namely $\mathcal{C}_R = \mathcal{C}_{in} = \mathcal{C}_{out}$ and $\mathcal{C}_M$. Clearly, $\overrightarrow{K}_h \in D$ implies that $\mathcal{C}_R \in D$ or $\mathcal{C}_M \in D$.

Let us show that $\tau_2 \to \infty$ implies $P(\mathcal{C}_R \in D) \to 1$. If $mp_-p_+ \to 0$ this follows by Lemma 4.1. Otherwise, using Lemma 3.1 we may assume $mp_-p_+ \to c \in (0; \infty]$. Consider another sequence of random digraphs $\{D'(k)\}$ where $D'(k) = D(n, m, (\omega m p_+)^{-1}, p_+)$ and $\omega = \ln n$. By Lemma 4.1 we have that $\mathcal{C}_R \in D'(k)$ whp, and by monotonicity $\mathcal{C}_R \in D$ whp.

Now suppose $\tau_2 = O(1)$. Then $\tau_1 = m^{-1/2}\tau_2 \to 0$ and $\mu(\mathcal{C}_M) = \tau_1^h \to 0$. This implies by (3) that any simple cover $\mathcal{C}$ of $\overrightarrow{K}_2$ such that $\mathcal{C} \neq \mathcal{C}_R$ must have $\mu(\mathcal{C}) \to 0$. By Lemma 4.1 and Lemma 3.4 for any proper diclique cover $\mathcal{C} \neq \mathcal{C}_R$ we have $P(\mathcal{C} \in D) \to 0$. Finally Remark 4.2 shows that $P(\mathcal{C}_R \in D) = \Omega(1)$ if $\tau_2 = \Theta(1)$ and so $\mathcal{C}_R$ is indeed the leading cover. $\square$

A Proofs of (1)-(8)

We will prove the estimates from Section 3.

Proof of (1) - (8) Similarly as in [13] we can represent D by two random n by m binary matrices R^- and R^+ where $R^-_{ij} = 1$ if and only if $j \in S^-(i)$ and $R^+_{ij} = 1$ if and only if $j \in S^+(i)$.

To prove (1) we will apply Lemma 2 from [13]. The probability of success p is the probability that a fixed key generates C on a fixed set $S \supseteq C^- \cup C^+$:

$$p = p_-^{|C^-|} p_+^{|C^+|} (1-p_-)^{|S|-|C^-|} (1-p_+)^{|S|-|C^+|}.$$

We have that $p \sim p_-^{|C^-|} p_+^{|C^+|}$ since $|S|$ is fixed and $p_-, p_+ \to 0$. Since C is proper and $mp_-p_+ \to 0$ we have that $mp \to 0$. By independence and the inclusion-exclusion principle (or by Lemma 2 of [13]) we have that $P(C) \sim mp \sim mp_-^{|C^-|} p_+^{|C^+|}$. Equation (2) follows from analogous reasoning as in the proof of Theorem 3 of [13]. For the random digraph D let N_C count the number of different attributes $w \in W$ that generate the diclique C (on the set $V(\mathcal{C})$). Suppose $\{C_1, C_2, \dots, C_t\}$ are all distinct dicliques in $\mathcal{C}$ where C_i has multiplicity a_i, $i = 1 \dots t$ and let $\{C_{t+1}, \dots, C_M\}$ be the set of all proper dicliques on $V(\mathcal{C})$ that are not in $\mathcal{C}$. Then by Lemma 1 of [13]:

$$\begin{aligned}
P(\mathcal{C}) &= P(N_{C_1} = a_1, \dots, N_{C_t} = a_t, N_{C_{t+1}} = 0, \dots, N_{C_M} = 0) \\
&\sim P(N_{C_1} \geq a_1, \dots, N_{C_t} \geq a_t, N_{C_{t+1}} = 0, \dots, N_{C_M} = 0) \\
&\sim \frac{P(C_1)^{a_1}}{a_1!} \frac{P(C_2)^{a_2}}{a_2!} \dots \frac{P(C_t)^{a_t}}{a_t!} \sim a_{\mathcal{C}} m^{|\mathcal{C}|} p_-^{\sum \mathcal{C}^-} p_+^{\sum \mathcal{C}^+}
\end{aligned}$$

since for $j > t$, $P(N_{C_j} = 0) = 1 - P(C_j) \to 1$. Now the equation (3) is immediate since $EX(\mathcal{C}) = \binom{n}{|V(\mathcal{C})|} P(\mathcal{C})$.

To see (5), recall that by Lemmas 1 and 2 of [13] we have for any diclique family $\mathcal{C}$:

$$P(\mathcal{C}) \sim \prod_{C \in \mathcal{C}} P(C) \sim K a_{\mathcal{C}} \prod_{C \in \hat{\mathcal{C}}} P(C)$$

where $K = \prod_{C \in \mathcal{C} \setminus \hat{\mathcal{C}}} P(C)$ is a constant.

The proof of the upper bounds (6) and (7) is much simpler: we sum the probability that D contains a fixed copy of $\mathcal{C}$ realised by a fixed tuple of attributes, $\prod_{C\in\mathcal{C}} \tilde{P}(C)$, over all $\binom{m}{a_1,\dots,a_t}$ ways to pick a relevant tuple, and, for the bound (7), over all $\binom{n}{|V(\mathcal{C})|}$ sets of V of size $|V(C)|$. The estimate follows by the union bound. The estimate (8) follows similarly. $\square$

Acknowledgement I would like to thank Mindaugas Bloznelis for introducing me to this topic, useful remarks and for his help preparing the paper. I would also like to thank the anonymous reviewers for their comments which helped to improve this paper a lot.

References

[1] M. Behrisch, Component evolution in random intersection graphs, *The Electronic Journal of Combinatorics* **14** (2007).

[2] S. R. Blackburn and S. Gerke, Connectivity of the uniform random intersection graph, *Discr. Math.* **309** (2009) 5130–5140.

[3] M. Bloznelis, Degree distribution of a typical vertex in a general random intersection graph, *Lithuanian Mathematical Journal* **48** (2008) 38–45.

[4] M. Bloznelis, A random intersection digraph: Indegree and outdegree distributions, *Discr. Math.* **310** (2010) 2560–2566.

[5] M. Bloznelis, Component evolution in general random intersection graphs, *SIAM J. Discr. Math.* **24** (2010) 639–654.

[6] M. Deijfen and W. Kets, Random intersection graphs with tunable degree distribution and clustering, *Probab. Eng. Inf. Sci.* **23** (2009) 661-674.

[7] P. Erdős, A.W. Goodman and L. Pósa, The representation of a graph by set intersections. *Canad. J. Math.* **18** (1966) 106–112.

[8] L. Eschenauer and V.D. Gligor, A key-management scheme for distributed sensor networks, *Proceedings of the 9th ACM Conference on Computer and Communications Security* (2002) 41–47.

[9] J. Fill , E. Scheinerman , K. Singer-Cohen, Random intersection graphs when $m = \omega(n)$: an equivalence theorem relating the evolution of the $G(n, m, p)$ and $G(n, p)$ models, *Random Struct. Algorithms*, **16**, 156–176, 2000.

[10] E. Godehardt and J. Jaworski, Two models of random intersection graphs for classification. In: O. Optiz and M. Schwaiger, Editors, *Studies in Classification, Data Analysis and Knowledge Organization* **22**, Springer, Berlin (2003), 67–82.

[11] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, Wiley-Interscience, 2000.

[12] R. M. Haralick, The diclique representation and decomposition of binary relations, *J. ACM* **21** (1974) 356–366.

[13] M. Karoński, E.R. Scheinerman and K.B. Singer-Cohen, On Random Intersection Graphs: The Subgraph Problem, *Comb. Probab. Comput.* **8** (1999) 131–159.

[14] K. Rybarczyk, Equivalence of the random intersection graph and G(n,p), *Random Struct. Algorithms*, **38** (2011), 205–234.

[15] K. Rybarczyk, Diameter, connectivity, and phase transition of the uniform random intersection graph, *Discrete Mathematics*, **311** (2011) 1998–2019.

[16] K. Rybarczyk and D. Stark, Poisson approximation of the number of cliques in random intersection graphs, *J. Appl. Probab.*, **47** (2010), 826–840.

[17] K. Singer, Random Intersection Graphs, Ph.D. thesis, The Johns Hopkins University, 1995.

[18] D. Stark, The vertex degree distribution of random intersection graphs, *Random Struct. Algorithms* **24** (2004) 249–258.