# Randomness in graphs and the bunkbed conjecture

Erik Aas

LiTH - MAT - EX - - 2010 / 12 - - SE

# Randomness in graphs and the bunkbed conjecture

Department of Mathematics, Linköpings Universitet

**Erik Aas**

LiTH - MAT - EX - - 2010 / 12 - - SE

Examensarbete: **45 hp**

Level: **D**

Supervisor: **Svante Linusson**,
Department of Mathematics, Royal Institute of Technology

Examiner: **Bengt-Ove Turesson**,
Department of Mathematics, Linköpings Universitet

Linköping: **May 2010**

# Abstract

The bunkbed conjecture is an intuitively plausible, unproven conjecture concerning open paths in a random graph. We study this conjecture and problems naturally connected to it. In particular, an interesting connection between random graphs and random orientations of graphs is investigated.

# Acknowledgements

First of all, I would like to thank my supervisor Svante Linusson for suggesting most of the topics discussed here, for careful reading of earlier manuscripts, and more. Svante's ability to explain concepts in a clear and intutive manner has been particularly striking.

My examiner Bengt-Ove Turesson has been very helpful with all administrative aspects of this thesis. In fact, the same applies for my studies in general, too. Thanks!

Thanks to my opponent Jonas Strandman for providing helpful comments on earlier versions.

Though not in direct connection with this thesis, I have learned much on graph theory and on writing mathematics from Armen Asratian. This has been beneficial for me in general, not only in writing this thesis, and for this he owes much thanks.

For LaTeX templates and discussions I thank Jonas Karlsson.

# Contents

# Summary

All topics discussed will in some way be related to the bunkbed conjecture. In most cases this discussion is reduced to investigating how the events $\{x \to y\}$ and $\{y \to z\}$ and their probabilities are correlated, where $\{x \to y\}$ loosely means 'there is a path from $x$ to $y$', the specific definition varying slightly between the models considered.

- In Chapter 1 we compute the covariance between $\{x \to y\}$ and $\{y \to z\}$ in a random orientation of a complete bipartite graph, where $\{x \to y\}$ denotes the event that there is a directed path from $x$ to $y$ consistent with the orientation.
- Chapter 2 deals with the same covariance as above, now conditional on $\{z \to y\}$, in a randomly oriented random graph.
- The results in chapter 3 are concerned with an interesting similarity between percolation in undirected graphs and percolation in directed graphs
- Chapter 4, describing the bunkbed conjecture, is the main part of this thesis.
- Finally in Chapter 5, some topics related to the bunkbed conjecture are discussed.

# Chapter 1

# Directed paths in random orientations of $K_{m,n}$

Let $x$, $y$ and $z$ be three distinct vertices in the complete bipartite graph $K_{m,n}$ with partition $(X, Y)$, where $|X| = m$, $|Y| = n$. By giving $K_{m,n}$ a uniformly chosen random orientation - that is, declaring each edge to be directed in either of the two possible ways with probability 1/2, independently of all other edges - we obtain a probability space with measure $P = P_{m,n}$. For any two vertices $x$, $y$ of $K_{m,n}$, say that $x$ *reaches* $y$, denoted by $x \to y$, if there is a directed path from $x$ to $y$ consistent with the chosen orientation. The goal of this chapter is to determine the asymptotic (as $m$ and $n \to \infty$ in a certain fashion; see below) covariance between the events $x \to y$ and $y \to z$. That is, we are interested in the behaviour of

$$P(x \to y \to z) - P(x \to y)P(y \to z)$$

as $n$ tends to infinity.

For any two events $A$, $B$ in any probability space, $P(A)P(B) - P(A \cap B) = P(A^c)P(B^c) - P(A^c \cap B^c)$. Hence one may consider

$$P(x \not\to y \not\to z) - P(x \not\to y)P(y \not\to z)$$

instead, which turns out to be easier.

To achieve this we need to bound the probability of various events related to the problem, explained below.

Clearly the covariance will depend on how $x$, $y$ and $z$ are situated in the graph.

From symmetry considerations, all possibilities are covered by the following three cases:

- $x, y, z \in X$,
- $x, y \in X$, $z \in Y$, and
- $x, z \in X$, $y \in Y$.

If $S$ and $T$ are two disjoint sets of vertices in $K_{m,n}$, an '$ST$-witness' is defined to be a vertex $u$ for which there is at least one edge from $S$ to $u$ and at least one edge from $u$ to $T$.

For a vertex $a$ in $K_{m,n}$, the set $O_a$ will loosely be defined as the set of vertices in $X' \cup Y'$ (thus in $X'$ if $a \in Y$ and in $Y'$ if $a \in X$) which can be reached in exactly one step from $a$, $X'$ and $Y'$ being defined separately in each section where this notation is used. We denote $|X'|$ and $|Y'|$ by $m'$ and $n'$. The set $I_a$ of vertices which reach $a$ in exactly one step is similarly defined.

## 1.1 Preliminary bounds

Knowing the following probabilities (modulo the obvious of symmetries of reversing all arrows and interchanging $X$ and $Y$) suffice to compute the covariance between the events $\{x \not\to z\}$ and $\{z \not\to y\}$:

- (i) $P(b \not\to a)$
- (ii) $P(d \not\to a)$
- (iii) $P(b \not\to d \not\to a)$
- (iv) $P(c \not\to b \not\to a)$
- (v) $P(d \not\to b \not\to a)$

Here, $a$, $b$, $c \in X$ and $d \in Y$ are four distinct vertices. In giving asymptotic bounds of these probabilties, we will restrict ourselves to the case $m = \lfloor \beta n \rfloor$, for some constant $\beta > 0$. However this turns out to be a rather general case, as the computed covariance will only depend on whether $\beta < 1$, $\beta = 1$, or $\beta > 1$. Thus, throughout we assume $\beta$ to be some given positive constant and $m = \lfloor \beta n \rfloor$.

Estimating these probabilities will be a lot of repetitive work. The following inequality will be used several times: if $s$, $t \geq \alpha$, then $st \geq \alpha s + \alpha t - \alpha^2$.

Below, when summing over subsets of nodes denoted by upper case letters, the sizes of these sets will often be denoted by the corresponding lower case letters.

**(i)** $P(b \not\to a)$

**Lemma 1.1.**

$$P(b \not\to a) \sim 2\left(\frac{1}{2}\right)^n.$$

*Proof.* Let $X' = X - \{a, b\}$, $Y' = Y$.

A lower bound is given by $P(b \not\to a) \geq P(\{\text{there is no edge directed away from } b\} \cup \{\text{there is no edge directed towards } a\}) = 2\left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^{2n}$, by inclusion-exclusion.

By calculating the probability that there is no path from $b$ to $a$ of length at most 4, we get the following upper bound: $P(b \not\to a) = \sum_{S,T \subseteq Y} P(b \not\to a | O_b = S, I_a = T) P(O_b = S, I_a = T) \leq \left(\frac{1}{2}\right)^{2n} \sum_{S,T \subseteq Y : S \cap T = \emptyset} P(\text{no } x \in X' \text{ is an } ST\text{-witness}) = \left(\frac{1}{2}\right)^{2n} \sum_{s=0}^{n} \sum_{t=0}^{n-s} \binom{n}{s}\binom{n-s}{t} \left(\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t - \left(\frac{1}{2}\right)^{s+t}\right)^{m-2}$.

Note that the partial sum corresponding to $st = 0$ is equal to the lower bound. We now show that the other terms sum to $o(\left(\frac{1}{2}\right)^n)$. Split the remaining sum into the following four parts: $S_1$: $s$, $t \geq \alpha$; $S_2$: $1 \leq s \leq \alpha \leq t$; $S_3$: $1 \leq t \leq \alpha \leq s$; $S_4 : 1 \leq s, t \leq \alpha$.

Note that in the $S_1$ case, $\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t - \left(\frac{1}{2}\right)^{s+t} \le \left(\frac{1}{2}\right)^{\alpha-1}$. Hence $S_1 \le \left(\frac{1}{2}\right)^{2n} \sum_{s=\alpha}^{n} \binom{n}{s} \sum_{t=\alpha}^{n-s} \binom{n-s}{t} \left(\frac{1}{2}\right)^{(\alpha-1)(m-2)} = \left(\frac{1}{2}\right)^{2n+(\alpha-1)(m-2)} 3^n = o(\left(\frac{1}{2}\right)^{(\alpha-1)(m-2)}) = o(\left(\frac{1}{2}\right)^n)$, the last equality holding when choosing $\alpha$ large enough.

$S_2 \le \left(\frac{1}{2}\right)^{2n} \sum_{s=1}^{\alpha} \binom{n}{s} \sum_{t=\alpha}^{n-s} \binom{n-s}{t} \left(\left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^t\right)^{m-2} \le \left(\frac{1}{2}\right)^{2n} \sum_{s=1}^{\alpha} n^\alpha \sum_{t=0}^{n} \binom{n}{t} \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^\alpha\right)^{m-2} \le \left(\frac{1}{2}\right)^n \alpha n^\alpha \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^\alpha\right)^{m-2} = o(\left(\frac{1}{2}\right)^n)$, if $\alpha > 0$.

By symmetry, we may choose $\alpha$ possbily even larger so that $S_3 = o(\left(\frac{1}{2}\right)^{2n})$ holds.

Clearly, $S_4 = o(\left(\frac{1}{2}\right)^n)$.

Hence $P(b \not\to a) - 2\left(\frac{1}{2}\right)^n \le S_1 + S_2 + S_3 + S_4 = o(\left(\frac{1}{2}\right)^{2n})$. □

**(ii)** $P(d \not\to a)$

**Lemma 1.2.** *$P(d \not\to a) \sim \left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n$. (This is $\sim \left(\frac{1}{2}\right)^m$ for $\beta < 1$, $\sim 2\left(\frac{1}{2}\right)^n$ for $\beta = 1$, and $\sim \left(\frac{1}{2}\right)^n$ for $\beta > 1$.)*

*Proof.* Let $X' = X - \{a\}$, $Y' = Y - \{d\}$. The probability is bounded from below by $P(d \not\to a) \ge P(\text{no edge leaves } d \text{ or no edge enters } a) \ge \left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n - \left(\frac{1}{2}\right)^{m+n-1}$.

For the upper bound, we calculate the probability that there is no path from $d$ to $a$ of length at most 3: $P(d \not\to a) = \sum_{S\subseteq Y', T\subseteq X'} P(a \not\to d | I_a = S, O_d = T)P(I_a = S, O_d = T) \le \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} P(\text{no edge from } S \text{ to } T) = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st}$. The partial sum with $st = 0$ equals the lower bound. We now show that the remaining terms sum to $o(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n)$ by splitting their sum into the following cases: $S_1$: $s, t \ge \alpha$, $S_2$: $1 \le s \le \alpha$, and $S_3$: $1 \le t \le \alpha$.

Using $s, t \ge \alpha \Rightarrow st \ge \alpha s + \alpha t - \alpha^2$, $S_1 = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st} \le \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{\alpha s+\alpha t-\alpha^2} \le \left(\frac{1}{2}\right)^{m+n-1} 2^{\alpha^2} \sum_{s=0}^{n-1} \binom{n-1}{s} \left(\frac{1}{2}\right)^{\alpha s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{\alpha t} \le \left(\frac{1}{2}\right)^{m+n-1} 2^{\alpha^2} \left(1 + \left(\frac{1}{2}\right)^\alpha\right)^{m+n-2} = o(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n)$, choosing $\alpha$ large enough. Similarly, $S_2 = \left(\frac{1}{2}\right)^{m+n-1} \sum_{s=1}^{\alpha} \binom{n-1}{s} \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^{st} \le \left(\frac{1}{2}\right)^{m+n-1} (n-1)^\alpha \alpha \sum_{t=0}^{m-1} \binom{m-1}{t} \left(\frac{1}{2}\right)^t = \alpha(n-1)^\alpha \left(\frac{1}{2}\right)^n \left(\frac{3}{4}\right)^{m-1} = o(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n)$.

A similar argument shows $S_3 = o(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n)$.

Hence $P(d \not\to a) - (\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n) \le S_1 + S_2 + S_3 = o(\left(\frac{1}{2}\right)^m + \left(\frac{1}{2}\right)^n)$. □

**(iii)** $P(b \not\to d \not\to a)$

**Lemma 1.3.** *$P(b \not\to d \not\to a) \sim 2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}$*

*Proof.* Let $X' = X - \{a, b\}$, $Y' = Y - \{d\}$. For the lower bound, we calculate the probability $P(\{$ the edge between $b$ and $d$, and the edge between $d$ and $a$, form a directed path from $a$ to $b\} \cap (\{O_b = O_d = \emptyset\} \cup \{O_b = I_a = \emptyset\} \cup \{I_d = I_a = \emptyset\}))$, which is $2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n} - \left(\frac{1}{2}\right)^{m+2n-3}$, by inclusion-exclusion; $P(b \not\to d \not\to a) \ge 2\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n} - \left(\frac{1}{2}\right)^{m+2n-3}$.

To get a working upper bound, it is sufficient to calculate the probability of there being no path from $b$ to $d$ or from $d$ to $a$, either of length at most 3. Conditioning on $I_a = S, O_b = T, I_d = U, O_d = V$, there may be no edge from $T$

to $U$, nor from $V$ to $S$. The edges $\{b, d\}$ and $\{a, d\}$ form a directed path from $a$ to $b$. This implies that $S$ and $T$ must be disjoint.

Hence

$$P(b \not\to d \not\to a)$$

$$= \sum_{S,T,U,V} P(b \not\to d \not\to a | I_a = S, O_b = T, I_d = U, O_d = V) P(I_a = S, O_b = T, I_d = U, O_d = V)$$

$$\leq \left(\frac{1}{2}\right)^{m+2n-2} \sum_{s=0}^{n-1} \binom{n-1}{s} \sum_{t=0}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \binom{m-2}{u} \left(\frac{1}{2}\right)^{tu+s(m-2-u)}.$$

The sum of the terms for which $s = t = 0$, $t = u = 0$ or $t = u - (m-2) = 0$ equals the lower bound. The remaining sum is split into the following cases: $S_1$: $s, t \geq \alpha$, $S_2$: $1 \leq t \leq \alpha \leq s$, $S_3 : 1 \leq s \leq \alpha \leq t$, and $S_4 : 1 \leq s, t \leq \alpha$.

$S_1 \leq 2^{2\alpha^2} \left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \sum_{t=\alpha}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \left(\frac{1}{2}\right)^{\alpha(m-2)} = \left(\frac{1}{2}\right)^{2n+m-2} 3^{n-1} \left(\frac{1}{2}\right)^{(\alpha-1)(m-2)} = o(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n})$.

$S_2 \leq \left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=1}^{\alpha} \binom{n-1}{s} \sum_{t=\alpha}^{n-1-s} \binom{n-1-s}{t} \sum_{u=0}^{m-2} \left(\frac{1}{2}\right)^{tu+s(m-2-u)} \leq \left(\frac{1}{2}\right)^{2n+m-2} (n-1)^\alpha \alpha \sum_{t=\alpha}^{n-1} \binom{n-1}{t} \sum_{u=0}^{m-2} \binom{m-2}{u} \left(\frac{1}{2}\right)^{\alpha u+m-2-u} = \left(\frac{1}{2}\right)^{2(m-2+n)} (n-1)^\alpha \alpha 2^{n-1} \left(1 + \left(\frac{1}{2}\right)^{\alpha-1}\right)^{m-2} = o(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n})$.

By symmetry with $S_2$, we deduce $S_3 = o(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n})$.

$S_4 \leq \left(\frac{1}{2}\right)^{2n+m-2} (n-1)^{2\alpha} \alpha^2 = o(\left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n})$.

We conclude that

$$P(b \not\to d \not\to a) \sim 2 \left(\frac{1}{2}\right)^{m+n-1} + \left(\frac{1}{2}\right)^{2n}.$$

□

**(iv)** $c \not\to b \not\to a$

**Lemma 1.4.**

$$P(c \not\to b \not\to a) \sim 3 \left(\frac{1}{2}\right)^{2n}$$

*Proof.* For the lower bound, note that $P(c \not\to b \not\to a) \geq P(\{O_c = O_b = \emptyset\} \cup \{O_c = I_a = \emptyset\} \cup \{I_b = I_a = \emptyset\}) \geq 3 \left(\frac{1}{2}\right)^{2n} - 2 \left(\frac{1}{2}\right)^{3n}$.

For the upper bound, we will sum over $U \subseteq Y'$, $V = Y' - U$, $S \subseteq U$, and $T \subseteq V$. When doing so, an expression for the probability that a given vertex $x' \in X'$ is not a $TU$-witness and not a $VS$-witness is needed. The probability of the complementary event is the probability of $x'$ being a $TU$-witness or a $VS$-witness. The separate probabilities for these last two events are $P(x'$ is a $TU$-witness$) = \left(1 - \left(\frac{1}{2}\right)^{|T|}\right)\left(1 - \left(\frac{1}{2}\right)^{|U|}\right)$ and $P(x'$ is an $SV$-witness$) = \left(1 - \left(\frac{1}{2}\right)^{|S|}\right)\left(1 - \left(\frac{1}{2}\right)^{|V|}\right)$. The probability of their intersection is $\left(1 - \left(\frac{1}{2}\right)^{|S|}\right)\left(1 - \left(\frac{1}{2}\right)^{|T|}\right)$, using $S \subseteq U$, $T \subseteq V$, By inclusion-exclusion, we get

$P(x \in X'$ is not a $TU$-witness, nor a $VS$-witness$) = 1 - \left(1 - \left(\frac{1}{2}\right)^{|T|}\right)\left(1 - \left(\frac{1}{2}\right)^{|U|}\right) - \left(1 - \left(\frac{1}{2}\right)^{|S|}\right)\left(1 - \left(\frac{1}{2}\right)^{|V|}\right) + \left(1 - \left(\frac{1}{2}\right)^{|S|}\right)\left(1 - \left(\frac{1}{2}\right)^{|T|}\right)$, which simplifies to $\left(\frac{1}{2}\right)^{|U|} + \left(\frac{1}{2}\right)^{|V|} - \left(\frac{1}{2}\right)^{|T|+|U|} - \left(\frac{1}{2}\right)^{|S|+|V|} + \left(\frac{1}{2}\right)^{|S|+|T|}$.

$$P(c \not\to b \not\to a)$$

$$= \sum_{S,T,U,V} P(c \not\to b \not\to a | I_a = S, O_c = T, I_b = U, O_b = V) P(I_a = S, O_c = T, I_b = U, O_b = V)$$

$$\leq \sum_{S,T,U,V} P(\text{no } x \in X' \text{ is a } TU\text{-witness, nor a } VS\text{-witness}) \left(\frac{1}{2}\right)^{3n}$$

$$= \left(\frac{1}{2}\right)^{3n} \sum_{u=0,u+v=n}^{n} \binom{n}{u} \sum_{s=0}^{u} \binom{u}{s} \sum_{t=0}^{v} \binom{v}{t} \left(\left(\frac{1}{2}\right)^{u} + \left(\frac{1}{2}\right)^{v} - \left(\frac{1}{2}\right)^{t+u} - \left(\frac{1}{2}\right)^{s+v} + \left(\frac{1}{2}\right)^{s+t}\right)^{m-3}.$$

The sum of the terms with $s = t = 0$ or $u = 0$ or $v = 0$ equals the lower bound. The other terms sum to $o(\left(\frac{1}{2}\right)^{2n})$, as we now turn to show. Since $(u, t)$ and $(v, s)$ are interchangeable, we need only consider the following cases: $S_1$: $s$, $t \geq \alpha$; $S_2$: $1 \leq s \leq \alpha \leq t, u$; $S_3$: $1 \leq s \leq \alpha \leq t$, $1 \leq u \leq \alpha$; $S_4$: $1 \leq s, t \leq \alpha$.

$$S_1 \leq \left(\frac{1}{2}\right)^{3n} \sum_{u=\alpha,u+v=n}^{n-\alpha} \binom{n}{u} \sum_{s=\alpha}^{u} \binom{u}{s} \sum_{t=0}^{v} \binom{v}{t} \left(\left(\frac{1}{2}\right)^{u} + \left(\frac{1}{2}\right)^{v} - \left(\frac{1}{2}\right)^{t+u} - \left(\frac{1}{2}\right)^{s+v} + \left(\frac{1}{2}\right)^{s+t}\right)^{m-3}$$

$$\leq \left(\frac{1}{2}\right)^{3n} \sum_{u=0,u+v=n}^{n} \binom{n}{u} \sum_{s=0}^{u} \binom{u}{s} \sum_{t=0}^{v} \binom{v}{t} \left(\frac{1}{2}\right)^{(\alpha-2)(m-3)} = O(\left(\frac{1}{2}\right)^{(\beta(\alpha-2)+1)n}),$$

which is $o(\left(\frac{1}{2}\right)^{2n})$ when choosing $\alpha$ large enough (e.g. $\alpha > 2(1 + 1/\beta)$).

Note that $t \geq \alpha \Rightarrow v \geq \alpha$.

$S_2 \leq \left(\frac{1}{2}\right)^{3n} \sum_{u=\alpha}^{n-\alpha} \binom{n}{u} \sum_{s=1}^{\alpha} n^{\alpha} \sum_{t=\alpha}^{v} \binom{v}{t} \left(3 \left(\frac{1}{2}\right)^{\alpha}\right)^{m-3} = o(\left(\frac{1}{2}\right)^{2n})$ for large enough $\alpha$.

$S_3 \leq \left(\frac{1}{2}\right)^{3n} \sum_{u=1}^{\alpha} n^{2\alpha} \sum_{s=1}^{\alpha} \sum_{t=\alpha}^{v} \binom{v}{t} \left(\left(\frac{1}{2}\right)^{\alpha} + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-3} = o(\left(\frac{1}{2}\right)^{2n})$ for large enough $\alpha$.

$S_4 \leq \left(\frac{1}{2}\right)^{3n} \alpha^3 n^{3\alpha} = o(\left(\frac{1}{2}\right)^{2n})$.

Consequently $P(c \not\to b \not\to a) - 3\left(\frac{1}{2}\right)^{2n} = o(\left(\frac{1}{2}\right)^{2n})$. □

**(v)** $d \not\to b \not\to a$

**Lemma 1.5.** *$P(d \not\to b \not\to a) \sim \left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}$.*

*Proof.* Let $X' = X - \{a, b\}$, $Y' = Y - \{d\}$.

As before, we have the simple lower bound: $P(d \not\to b \not\to a) \geq P(\{O_d = O_b = \emptyset\} \cup \{O_d = I_a = \emptyset\} \cup \{I_b = I_a = \emptyset\}) \geq \left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n} - \left(\frac{1}{2}\right)^{m+2n-3}$.

We bound the probability from above by the probability of there being no path from $d$ to $b$ or from $b$ to $a$ of length at most 3 or 4 respectively. The edges $\{a, d\}$ and $\{b, d\}$ are both directed towards $d$. Condition on $O_d = T$, $I_b = U$, $I_a = S$, $O_b = V$. No edge is directed from $T$ to $U$, and $S \subseteq U$. These conditions imply that no $x \in S$ is a $VU$-witness. In addition we forbid any $x \in X' - T$ to be a $VS$-witness. The events '$x$ is a $VS$-witness' are independent for $x \in X' - T$ and independent of the other necessary events just stated. We obtain

$$P(d \not\to b \not\to a)$$

$$= \sum_{S,T,U,V} P(d \not\to b \not\to a | O_d = T, I_b = U, O_b = V, I_a = S) P(O_d = S, I_b = T, I_a = U, I_a = S)$$

$$\leq \left(\frac{1}{2}\right)^{2n+m-2} \sum_{t=0}^{m-2} \binom{m-2}{t} \sum_{u=0}^{n-1} \binom{n-1}{u} \sum_{s=0}^{u} \binom{u}{s} \left(\frac{1}{2}\right)^{st} \left(\left(\frac{1}{2}\right)^{s} + \left(\frac{1}{2}\right)^{v} - \left(\frac{1}{2}\right)^{s+v}\right)^{m-2-t}$$

$$= \left(\frac{1}{2}\right)^{2n+m-2} \sum_{u=0}^{n-1} \sum_{s=0}^{u} \binom{n-1}{u} \binom{u}{s} \left(\left(\frac{1}{2}\right)^{u} + \left(\frac{1}{2}\right)^{v} + \left(\frac{1}{2}\right)^{s} - \left(\frac{1}{2}\right)^{s+v}\right)^{m-2}.$$

For $s = 0$ we obtain the following sum: $\left(\frac{1}{2}\right)^{2n+m-2} \sum_{u=0}^{n-1} \binom{n-1}{u} \left(\left(\frac{1}{2}\right)^{u} + 1\right)^{m-2} = \left(\frac{1}{2}\right)^{2n+m-2} \left(2^{m-2} + (n-1)^{\alpha} \alpha \left(\frac{3}{2}\right)^{m-2} + \left(1 + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2}\right) = \left(\frac{1}{2}\right)^{2n} + (n-1)^{\alpha} \alpha \left(\frac{1}{2}\right)^{2n} \left(\frac{3}{2}\right)^{m-2} + \left(\frac{1}{2}\right)^{2n+m-2} \left(1 + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2} = \left(\frac{1}{2}\right)^{2n} + o(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n})$.

For $v = 0 = n-1-u$: $\left(\frac{1}{2}\right)^{2n+m-2} \sum_{s=0}^{n-1} \binom{n-1}{s} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = \left(\frac{1}{2}\right)^{m+n-2} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = \Theta(\left(\frac{1}{2}\right)^{m+n-1})$, as $\lim_{n\to\infty} \left(1 + \left(\frac{1}{2}\right)^{n-1}\right)^{m-2} = 1$.

Split the remaining sum into the following cases: $S_1 : 1 \leq u \leq \alpha$, $S_2 : 1 \leq s \leq \alpha \leq u, v$, $S_3 : \alpha \leq s, u, v$, $S_4 : 1 \leq s, v \leq \alpha$, $S_5 : 1 \leq v \leq \alpha \leq s$.

Clearly, $S_1 = o(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n})$.

$S_2 \leq n^{\alpha} \left(\frac{1}{2}\right)^{m+2n-2} \sum_{u=\alpha}^{n-\alpha} \binom{n-1}{u} \left(2 \left(\frac{1}{2}\right)^{\alpha} + \left(\frac{1}{2}\right)^{s}\right)^{m-2} = o(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n})$, for $\alpha > 21$, which is easily seen by considering $\beta \leq 1$ and $\beta > 1$ separately.

$S_3 \leq \left(\frac{1}{2}\right)^{m+2n-2} \sum_{u=\alpha}^{n-\alpha} \binom{n-1}{u} \sum_{s=\alpha}^{u} \binom{u}{s} \left(3 \cdot \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2} \leq \left(\frac{1}{2}\right)^{m+2n-2} \left(\frac{3}{2^{\alpha}}\right)^{m-2}$. $3^{n-1} = o(\left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n})$.

$S_4 = o(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n})$, since $S_4$ is the sum of a constant ($\alpha^2$) number of $o(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n})$ terms.

$S_5 \leq \alpha n^{\alpha} \left(\frac{1}{2}\right)^{m+2n-2} \sum_{s=\alpha}^{n-1} \binom{n-1}{s} \left(\left(\frac{1}{2}\right)^{\alpha} + \left(\frac{1}{2}\right)^{v} + \left(\frac{1}{2}\right)^{\alpha}\right)^{m-2} = \alpha n^{\alpha} \left(\frac{1}{2}\right)^{m+n} \left(\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{\alpha-1}\right)^{m-2} = o(\left(\frac{1}{2}\right)^{m+2n-2} + \left(\frac{1}{2}\right)^{2n})$ for $\alpha > 2$.

Finally,

$$P(d \not\to b \not\to a) \sim \left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n}.$$

□

## 1.2 Results

A good measure of the covariance between $\{x \not\leftrightarrow y\}$ and $\{y \not\leftrightarrow z\}$ is the *relative covariance*,

$$\frac{P(x \not\leftrightarrow y \not\leftrightarrow z) - P(x \not\leftrightarrow y)P(y \not\leftrightarrow z)}{P(x \not\leftrightarrow y \not\leftrightarrow z)}.$$

Note that this expression can be written as

$$1 - \frac{P(x \not\leftrightarrow y)P(y \not\leftrightarrow z)}{P(x \not\leftrightarrow y \not\leftrightarrow z)} = 1 - \frac{P(x \not\leftrightarrow y)}{P(x \not\leftrightarrow y | y \not\leftrightarrow z)}.$$

The results above can be used to compute the limit of the relative covariance between $\{x \not\leftrightarrow y\}$ and $\{y \not\leftrightarrow z\}$,

$$\lim_{n\to\infty} \frac{P(x \not\leftrightarrow y \not\leftrightarrow z) - P(x \not\leftrightarrow y)P(y \not\leftrightarrow z)}{P(x \not\leftrightarrow y \not\leftrightarrow z)}. \tag{1.1}$$

As an example, suppose $x, y \in X$, $z \in Y$ and $\beta < 1$.
Then

$$P(x \not\leftrightarrow y \not\leftrightarrow z) \sim \left(\frac{1}{2}\right)^{m+n-2} + \left(\frac{1}{2}\right)^{2n} \sim \left(\frac{1}{2}\right)^{2n},$$

the expression for $P(d \not\leftrightarrow b \not\leftrightarrow a)$.
Similarly

$$P(x \not\leftrightarrow y) \sim 2\left(\frac{1}{2}\right)^{n},$$

and

$$P(y \not\leftrightarrow z) \sim \left(\frac{1}{2}\right)^{m} + \left(\frac{1}{2}\right)^{n} \sim \left(\frac{1}{2}\right)^{m}.$$

Putting these results together we get

$$\lim_{n\to\infty} \frac{P(x \not\leftrightarrow y \not\leftrightarrow z) - P(x \not\leftrightarrow y)P(y \not\leftrightarrow z)}{P(x \not\leftrightarrow y \not\leftrightarrow z)} = \frac{1}{2},$$

for this choice of $x$, $y$, $z$, and $\beta$.
The other possible values of (1.1) are summarized below. The result of our example computation is underlined.

| $X$ | $Y$ | $\beta < 1$ | $m = n$ | $\beta > 1$ |
|---|---|---|---|---|
| $x, y, z$ | | -1/3 | -1/3 | -1/3 |
| $x, y$ | $z$ | <u>1/2</u> | 1/5 | -1 |
| $x, z$ | $y$ | 1 | 1/5 | 0 |

From these results, one might suspect $\{c \not\leftrightarrow b\}$ and $\{b \not\leftrightarrow a\}$ to be negatively correlated in any complete bipartite graph. However, numerical results suggest that when $n$ is much greater than $m$, the correlation is positive and can be arbitrarily close to 1 for arbitrarily large values of $m$ and $n$ (that is, of $\min(m, n)$). For example, the relative covariance in the case $m = 3$, $n = 31$, $x, y, z \in X$ is about 99.8%.

When $m \sim n$, the limit of the relative covariance is not given by the table above, except in the $m = n$ case. It is, however, not difficult to see that the value in the first row in the table actually is $-1/3$ even in the $m \sim n$ case.

**Another model**

Let $p \in [0,1]$ and let an edge $(x, y)$ with $x \in X$, $y \in Y$ be directed this way with probability $p$ and the other way with probability $1-p$. How do the results change?

## 1.3 Exact recursions

We will now derive formulas allowing for efficient calculation of the probabilities estimated above (as functions of $m$ and $n$; here we assume $m$ and $n$ to be free variables).

$P(a \not\to b)$**,** $P(a \not\to d)$

Let $X' = X - \{a, b\}$, $Y' = Y - \{d\}$.

Note that for arbitrary $x$, $y$, $P(x \not\to y) = P(y \not\to x)$ (reverse all arrows) and that the probabilities in the title may be written as $P(K \not\to a)$ for $K = \{b\}$ and $K = \{d\}$. Generalizing this to aribitrary sets $K \subseteq X$ or $K \subseteq Y$ we obtain functions of $K$, $m$, $n$ for which two simple recursion formulas hold.

As a notational convience, we will assume that $K \subseteq X'$ and $L \subseteq Y'$, $L$ playing a similar role to that of $K$. Now, let $f_X(m,n,K) = P(K \not\to a \text{ in } K_{m,n}) = P_{m,n}(K \not\to a)$, and $f_Y(m,n,L) = P_{m,n}(L \not\to a)$. Note that $f_X(m,n,K)$ and $f_X(m,n,L)$ depend only (apart from $m$ and $n$) on the cardinalities $|K|$ and $|L|$ of $K$ and $L$. Define $f_X(m,n,|K|)$ and $f_Y(m,n,|L|)$ to be these common values. We will often use $k$ and $l$ to denote $|K|$ and $|L|$ respectively. Hence, using the notation just introduced, we have $f_X(m,n,1) = P_{m,n}(a \not\to b)$ and $f_Y(m,n,1) = P_{m,n}(a \not\to d)$.

$P(O_K = L) = \left(1 - \left(\frac{1}{2}\right)^k\right)^l \left(\frac{1}{2}\right)^{(n-l)k} = \frac{(2^k-1)^l}{2^{nk}}$. Similarly, $P(O_L = K) = \frac{(2^l-k)^l}{2^{(m-k)l}}$. Now condition on $O_K = L$. Observe that $P_{m,n}(K \not\to a | O_K = L) = P_{m-k,n}(L \not\to a))$, since any $K \to a$-path passes through $L$ (thinking of $K_{m-k,n}$ as $K_{m,n} - K$). Similarly $P_{m,n}(L \not\to a | O_L = K) = P_{m,n-l}(K \not\to a) \left(\frac{1}{2}\right)^l$, the extra factor due to (cumbersome notation and) the fact $L \not\to a \Rightarrow$ no arrow from $L$ is directed towards $a$. Using this, one has

$$f_X(m,n,k) = f_X(m,n,K) = \sum_{L \subseteq Y'} P_{m,n}(K \not\to a | O_K = L) P(O_K = L)$$

$$= \sum_{l=0}^{n} \binom{n}{l} \frac{(2^k-1)^l}{2^{nk}} f_Y(m-k, n, l),$$

and

$$f_Y(m,n,l) = f_Y(m,n,L) = \sum_{K \subseteq X'} P_{m,n}(L \not\to a | O_L = K) P(O_L = K)$$

$$= \sum_{k=0}^{m-1} \binom{m-1}{l} \frac{(2^k-1)^l}{2^{mk}} f_X(m, n-l, k).$$

Observe that by definition, $f_X(m,n,0) = f_Y(m,n,0) = 1$ for all $m$, $n$.

**$P(d \not\to b \not\to a)$ and $P(c \not\to b \not\to a)$**

Let $g_X(m,n,K) = g_X(m,n,k) = P_{m,n}(K \not\to b \not\to a)$ and $g_Y(m,n,L) = g_Y(m,n,l) = P_{m,n}(L \not\to b \not\to a)$. Assume $O_K = L$. Then the following equality holds.

$$P_{m,n}(K \not\to b \not\to a | O_K = L) = P_{m-k,n}(L \not\to b \not\to a) \tag{1.2}$$

To prove (1.2), we argue that any orientation of $K_{m,n}$ whose restriction to $K_{m,n} - K$ satisfies $L \not\to b \not\to a$, satisfies $K \not\to b \not\to a$ (and the converse is obvious). To see this, observe that any $b \to a$-path in $K_{m,n}$ passes through $L$ if passing through $K$ and that all edges between $b$ and $L$ are directed away from $b$. In a similar (easier) vein,

$$P_{m,n}(L \not\to b \not\to a | O_L = K) = P_{m,n-l}(K \not\to b \not\to a).$$

We may now write down recursion formulas for $g_{\cdot}$:

$$g_X(m,n,k) = g_X(m,n,K) = P_{m,n}(K \not\to b \not\to a)$$

$$= \sum_{L \subseteq Y'} P_{m,n}(K \not\to b \not\to a | O_K = L) P_{m,n}(O_K = L)$$

$$= \sum_{l=0}^{n} \binom{n}{l} \frac{(2^k-1)^l}{2^{nk}} g_Y(m-k,n,l),$$

and

$$g_Y(m,n,l) = g_Y(m,n,L) = P_{m,n}(L \not\to b \not\to a)$$

$$= \sum_{K \subseteq X'} P_{m,n}(L \not\to b \not\to a | O_L = K) P(O_L = K) = \sum_{k=0}^{m-2} \binom{m-2}{k} \frac{(2^l-1)^k}{2^{lm}} g_X(m,n-l,k).$$

Note that $g_X(m,n,0) = g_Y(m,n,0) = f_X(m,n,1)$, since, by definition, the empty set reaches nothing.

**$P(b \not\to d \not\to a)$**

Proceeding as before, let $h_X(m,n,k) = P_{m,n}(K \not\to d \not\to a)$ and $h_Y(m,n,l) = P_{m,n}(L \not\to d \not\to a)$. Then

$$h_X(m,n,k) = \sum_{l=0}^{n-1} \binom{n-1}{l} \frac{(2^k-1)^l}{2^{nk}} h_Y(m-k,n,l),$$

and

$$h_Y(m,n,l) = \sum_{k=0}^{m-1} \binom{m-1}{k} \frac{(2^l-1)^k}{2^{mk}} h_X(m,n-l,k).$$

Additionally, $h_X(m,n,0) = h_Y(m,n,0) = f_Y(m,n,1)$.

# Random graphs

**Definition 1.1.** *Given a graph $G$ whose edges $e$ are labeled by real numbers $p_e$ lying in $[0, 1]$, we define the random graph associated to $G$, or simply the random graph $G$, to be the probability distribution on all subgraphs $\omega$ of $G$ satisfying*

$$P(\omega) = \prod_{e \in E(G): \omega_e = 1} p_e \prod_{e \in E(G): \omega_e = 0} (1 - p_e).$$

Here $\omega_e$ is 1 if the edge $e$ is in $\omega$ and 0 otherwise.

For convenience, we will say that $(G, \mathbf{p})$, or simply $G$, is a random graph, without risk of confusion.

We can generate an outcome $\omega$ of $G$ by deciding for each edge $e$ whether or not to include $e$ in $\omega$, the probability of including $e$ being determined by $p_e$, and these choices are made independently of each other. Note that any random graph $G$ is completely described by its underlying weighted graph $G$.

By choosing $G$ to be the complete graph on $n$ vertices and letting all edge labels equal some chosen $p \in [0, 1]$ we obtain the random graph $G(n, p)$, arguably the single most investigated random graph model (see, for example, [1]).

# Chapter 2

# A conditional correlation in G(n,p)

In this chapter we will study events in a random orientation of $G(n,p)$, that is, for each unordered pair of vertices $\{u, v\}$ (the total number of vertices is $n$),

- with probability $1-p$, there is no edge between $u$ and $v$,
- with probability $p/2$, there is a directed edge from $u$ to $v$,
- with probability $p/2$, there is a directed edge from $v$ to $u$,

these cases being mutually exclusive.

Let $a$, $b$, $c$ denote distinct vertices in a randomly oriented $G(n,p)$, $A = \{a \not\to c\}$, $B = \{c \not\to b\}$, $C = \{b \not\to a\}$, $x = p/2$, $y = 1-x$.

In [10] it is shown that the relative covariance satisfies

$$\frac{P(A\cap B) - P(A)P(B)}{P(A\cap B)} = 1 - \frac{P(A)P(B)}{P(A\cap B)} = \frac{2p-1}{3}. \tag{2.1}$$

We now show that when conditioning on $b \to a$, in our notation $C^c$, this limit persists.

**Theorem 2.1.** *Let $A$, $B$, $C$ be the events described above. Then*

$$\frac{P(a \not\to c \not\to b|b \to a) - P(a \not\to c|b \to a)P(c \not\to b|b \to a)}{P(a \not\to c \not\to b|b \to a)} = \frac{2p-1}{3}.$$

*Proof.* Using the notation introduced above, we may write the left hand side as

$$\frac{P(A\cap B|C^c) - P(A|C^c)P(B|C^c)}{P(A\cap B|C^c)} = 1 - \frac{P(A\cap C^c)P(B\cap C^c)}{P(C^c)P(A\cap B\cap C^c)}. \tag{2.2}$$

In [10] it is shown that $P(A) = P(B) = P(C) = O(y^n)$, $P(A\cap B) = P(A\cap C) = P(B\cap C) = O(y^{2n})$.

Observe that $P(A\cap C^c) = P(A) - P(A\cap C) = P(A) - P(A\cap B)$ and $P(A\cap B\cap C^c) = P(A\cap B) - P(A\cap B\cap C)$.

Suppose $P(A \cap B \cap C) = o(y^{2n})$. In this case, we can rewrite (2.2) as

$$1 - \frac{P(A)P(B)(1 - \frac{P(A\cap C)}{P(A)})(1 - \frac{P(B\cap C)}{P(B)})}{P(C^c)P(A \cap B)(1 - \frac{P(A\cap B\cap C)}{P(A\cap B)})},$$

which, by the remarks above, has the same limit as

$$1 - \frac{P(A)P(B)}{P(A \cap B)},$$

so what remains is to show that $P(A \cap B \cap C) = o(y^{2n}) = o(P(A \cap B))$ indeed holds.

Let $F_1 = \{O_a = O_c = \emptyset\}$, $F_2 = \{O_a = I_b = \emptyset\}$, $F_3 = \{I_c = I_b = \emptyset\}$, and $F = F_1 \cup F_2 \cup F_3$. Clearly, $F \subseteq A \cap B$. From [10],

$$P(F) = y^{2n-3}(3 - 2y^{2n-3}) \leq P(A \cap B) \leq y^{2n-3}(3 + o(1)).$$

Hence

$$0 \leq P(A \cap B) - P(F) \leq y^{2n-3}o(1) = o(y^{2n}),$$

and so we may write $A \cap B = (A \cap B \cap F) \cup (A \cap B \cap F^c) = F \cup (A \cap B \cap F^c)$, with $P(A \cap B \cap F^c) = o(y^{2n})$. Now in order to deduce $P(A \cap B \cap C) = o(y^{2n})$, we need to show that $P(F \cap C) = o(y^{2n})$. It is sufficient to show that $P(F_i \cap C) = o(y^{2n})$. It is intuitively clear that $P(F_2 \cap C) \leq P(F_1 \cap C) = P(F_3 \cap C)$. Since the three cases are all similar we prove only $P(C \cap F_1) = o(y^{2n})$. Note that $1 - p \leq 1 - p/2 = y$.

$$P(C \cap F_1) = \sum_{S,T \subseteq V - \{a,b,c\}: S \cap T = \emptyset} P(C \cap F_1 \cap \{I_a = S\} \cap \{O_b = T\})$$

$$= \sum_{s=0}^{n-3} \binom{n-3}{s} \sum_{t=0}^{n-3-s} \binom{n-3-s}{t} x^{s+t} y^{n-2-t} y (1-p)^{n-2-s} y^{n-1} P(\text{no edge from } S \text{ to } T)$$

$$= y^n \sum_{s=0}^{n-3} \binom{n-3}{s} \sum_{t=0}^{n-3} \binom{n-3}{t} x^{s+t} y^{2n-4-s-t} y^{st} = y^n O(y^{2n}) = o(y^{2n}),$$

the penultimate equality being more or less equivalent to lemma 2.1 in [10]. □

The original motivation of Theorem 2.1 was the conjecture of the events $\{x \to y\}$ and $\{y \to z\}$, conditional on $\{z \to y\}$, being nonnegatively correlated, in any random graph, which hence might admit a proof more conceptual than the computation given above. The theorem above shows that there are graphs for which this conjecture is false. Indeed, the graph in Figure 2 is a counterexample.

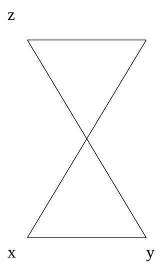


Figure 2.1: A counterexample

# Chapter 3

# Two equivalent models

In this chapter we will use the notion of mixed random graphs, random graphs having both undirected and directed edges. Note that this model is fundamentally different from the random orientation model considered earlier.

**Definition 3.1.** *A mixed random graph is a random graph in which each edge is either directed or undirected, and any edge e is present with some given probability* $p_e$*.*

The following theorem is given in [11].

**Theorem 3.1.** *Let* $G_1$ *be a mixed random graph and* $G_2$ *the result of picking an arbitrary undirected edge e in* $G_1$ *and turning it into two directed edges going in opposite directions, both new edges being open in* $G_2$ *with the same probability as e in* $G_1$*, independently of each other. Then*

$$P_{G_1}(x \to y) = P_{G_2}(x \to y).$$

*Proof.* Note that removing $e$ in $G_1$ or its two copies in $G_2$ gives the same graph, call it $G'$.

Let $\mathcal{H}_1 = \{x \to u, v \to y, x \not\to y \text{ in } G'\}$, $\mathcal{H}_2 = \{x \to v, u \to y, x \not\to y \text{ in } G'\}$, and $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$. Note that $\mathcal{H}_1 \cap \mathcal{H}_2 = \emptyset$. We regard $\mathcal{H}$ as an event in both $G_1$ and $G_2$ (with $P_{G_1}(\mathcal{H}) = P_{G_2}(\mathcal{H})$). The event $\mathcal{H}$ may intuitively be thought of as 'the state of the edge(s) between $u$ and $v$ determines whether $x \to y$ or $x \not\to y$, given the states of all other edges'.

Clearly $P_{G_1}(x \to y|\mathcal{H}^c) = P_{G_2}(x \to y|\mathcal{H}^c)$. Now, $P_{G_1}(x \to y|\mathcal{H}_1) = P_{G_2}(x \to y|\mathcal{H}_1) = p$, since $x \to y$, conditioned on $\mathcal{H}_1$, holds in $G_1$ or $G_2$ iff $\{u, v\}$ is open in $G_1$ or iff $(u, v)$ is open in $G_2$ respectively. We may use similar reasoning to show $P_{G_1}(x \to y|\mathcal{H}_2) = P_{G_2}(x \to y|\mathcal{H}_2)$. Collecting these facts gives $P_{G_1}(x \to y) = P_{G_1}(x \to y|\mathcal{H}^c)P_{G_1}(\mathcal{H}^c) + P_{G_1}(x \to y|\mathcal{H}_1)P_{G_1}(\mathcal{H}_1) + P_{G_1}(x \to y|\mathcal{H}_2)P_{G_1}(\mathcal{H}_2) = P_{G_2}(x \to y|\mathcal{H}^c)P_{G_2}(\mathcal{H}^c) + P_{G_2}(x \to y|\mathcal{H}_1)P_{G_2}(\mathcal{H}_1) + P_{G_2}(x \to y|\mathcal{H}_2)P_{G_2}(\mathcal{H}_2) = P_{G_2}(x \to y)$. □

**Remark 3.1.** *Repeated application of this theorem shows that the probability of* $\{x \to y\}$ *does not change when turning some set of undirected edges into directed edges or the other way around. In particular this probability is the same in the case when all edges are undirected as in the case when all edges are directed, a result proved in [3].*

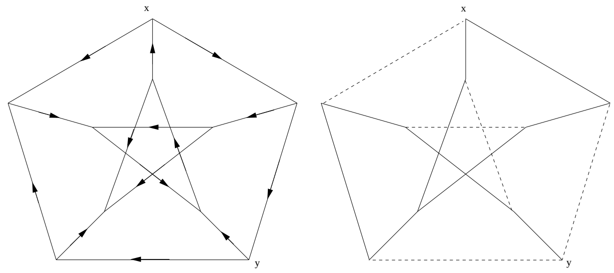


Figure 3.1: The graph on the left could be the result of giving the Petersen graph a uniformly chosen random orientation. The graph on the right could similarly be the outcome of a random graph process on the Petersen graph, where each edge is removed with probability 1/2 (drawn lines correspond to open edges, dashed lines to closed edges). As a consequence of Theorem 3.2, the probability of seeing a (directed) cycle is the same in the two graphs.

This theorem may be modified slightly to the notion of random orientation used in the previous chapter, as follows. Suppose instead $G_1$ has an undirected edge with parameter 1/2 and $G_2$ has a randomly (uniformly) oriented directed edge in the corresponding position. A similar proof shows that the $P(x \to y)$ is the same in $G_1$ and $G_2$.

In [3] a stronger result is proved, namely that the out-clusters in the two models have the same distribution. The argument above is easily modified to cover this result as well, by first proving that $P(\overrightarrow{C_x} \subseteq T) = P(\overrightarrow{C_x} \subseteq T)$ holds for any subset $T$ of vertices and from this deducing that $P(\overrightarrow{C_x} = T) = P(\overrightarrow{C_x} = T)$ for any $T$. To prove the first claim we could condition on each of the events $\mathcal{H}_S = \{x \to u, x \not\to u, \overrightarrow{C_u} = S\}$ and $\mathcal{H}'_S = \{x \to v, x \not\to u, \overrightarrow{C_v} = S\}$ for each subset $S$ of nodes and observe that $P(\overrightarrow{C_x} \subseteq T|\mathcal{H})$ is the same in $G_1$ and $G_2$, for any subset $T$ of nodes.

Another result in the same spirit (and which can be proved in a similar way) is the following (see [4]).

**Theorem 3.2.** *Let $G_1$ and $G_2$ be as in Theorem 3.1. Then*

$$P_{G_1}(\textit{There is a directed cycle}) = P_{G_2}(\textit{There is a directed cycle})$$

This result is illustrated in Figure 3.1.

# Chapter 4

# The bunkbed conjecture

## 4.1 Bunkbeds

Given an undirected random graph $G$ and a subset $T \subseteq V(G)$ of its vertices, define its associated *bunkbed graph* $B(G)$ *with transversal set* $T$, $B(G)$, as follows. Create two copies of $G$ (call one copy the lower layer, $G_0$, and the other one the upper layer, $G_1$), For each vertex $t \in T$, identify $t_0$ and $t_1$ and call the resulting vertex $t$ as well. These identifications potentially create parallel edges in the resulting graph. Remove any of the two edges in each such pair. The choice does not matter as the labels of two such edges necessarily are equal. Define $B(G)$ to be the resulting random graph. The formation of $B(G)$ from $G$ is illustrated for a particular graph $G$ in Figures 4.1-4.4. The set of vertices in the bunkbed graph created when identifying pairs of vertices coming from $T$ is naturally identified with $T$. We define $G_i' = G_i - T$ for $i = 1, 2$. If $x$ is a vertex of $G$, $x_0$ and $x_1$ will denote the two natural images of $x$ in $B(G)$. Similar notation will be used for edges. By convention, for vertices and edges $v$ and $e$ inside $T$ we have $v_0 = v_1$ and $e_0 = e_1$. For an edge $e_0$ or $e_1$ of the bunkbed graph, $p_e$ is the probability of either edge being open. Each edge is open independently of all other edges.

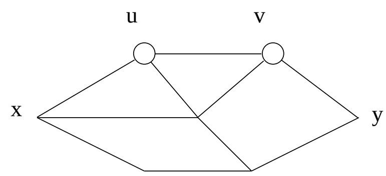


Figure 4.1: A graph with a distinguished set $T$ of vertices

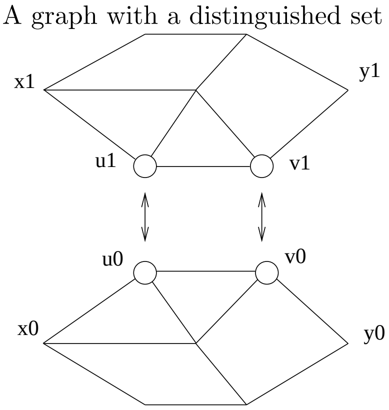


Figure 4.2: Two copies of the graph, indicating the identifications to be performed.

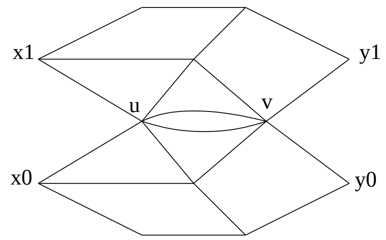


Figure 4.3: The result of identifying pairs of vertices coming from $T$.

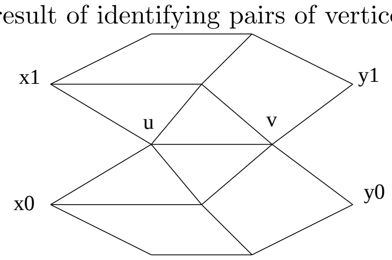


Figure 4.4: Here, one of the two parallel edges has been removed.

## 4.2 The bunkbed conjecture

The bunkbed conjecture (BBC): For any two vertices $x$, $y$ and subset of vertices $T$ of any random graph $G$,

$$P_{B(G)}(x_0 \leftrightarrow y_0) \geq P_{B(G)}(x_0 \leftrightarrow y_1).$$

This is intuitively clear. However the conjecture has resisted many (supposedly all) attempts of being proved rigourously. This conjecture was formulated by Kasteleyn in 1985, and since [8] it is known as the bunkbed conjecture.

The conjecture is clearly true if $x, y \notin T$ or $T = \emptyset$. These possibilities will be ignored in the sequel.

When $|T| = 1$, say $T = \{t\}$, the conjecture is a consequence of Harris' theorem on increasing events, viz. Clearly, $P(x_0 \leftrightarrow y_0 | x_0 \not\leftrightarrow t) \geq P(x_0 \leftrightarrow y_1 | x_0 \not\leftrightarrow t)$, since the right hand side is 0. Since $\{x_0 \leftrightarrow t\}$ and $\{t \leftrightarrow y_1\}$ are independent, $P(x_0 \leftrightarrow y_1 | x_0 \leftrightarrow t) = P(t \leftrightarrow y_1 | x_0 \leftrightarrow t) = P(t \leftrightarrow y_1)$. Hence $P(x_0 \leftrightarrow y_0 | x_0 \leftrightarrow t) = P(t \leftrightarrow y_0 | x_0 \leftrightarrow t) \geq P(t \leftrightarrow y_0) = P(t \leftrightarrow y_1)$, the inequality following from the following special case of Harris' theorem [13],

**Theorem 4.1.** *If $s, a, b, t$ are four (not necessarily distinct) vertices in some random graph $G$, the events $\{s \leftrightarrow t\}$ and $\{a \leftrightarrow b\}$ are positively correlated, that is*

$$P(s \leftrightarrow t, u \leftrightarrow v) \geq P(s \leftrightarrow t) P(u \leftrightarrow v),$$

*or, equivalently,*

$$P(s \leftrightarrow t | u \leftrightarrow v) \geq P(s \leftrightarrow t).$$

Hence $P(x_0 \leftrightarrow y_0) \geq P(x_0 \leftrightarrow y_1)$ in this case.

In [6] the following interesting generalisation of the theorem above is proved.

**Theorem 4.2.** *For any $s, a, b, t \in V(G)$, $G$ a random graph,*

$$P(s \leftrightarrow a, s \leftrightarrow b | s \not\leftrightarrow t) \geq P(s \leftrightarrow a | s \not\leftrightarrow t) P(s \leftrightarrow b | s \not\leftrightarrow t).$$

This statement, in contrast to Harris' theorem, is not intuitively clear (at least not to the present author).

The motivation of Theorem 4.2 was in fact the bunkbed conjecture. The theorem implies that $P(x_0 \leftrightarrow y_0 | s \not\leftrightarrow t) \geq P(x_0 \leftrightarrow y_1 | s \not\leftrightarrow t)$ when $T = \{s, t\}$ contains only two vertices. Hence what remains open in the case $|T| = 2$ is to show that $P(x_0 \leftrightarrow y_0 | s \leftrightarrow t) \geq P(x_0 \leftrightarrow y_1 | s \leftrightarrow t)$. It is interesting to note that, conditional on $\{s \leftrightarrow t\}$, the events $\{s \leftrightarrow a\}$ and $\{s \leftrightarrow b\}$, are not positively correlated in general. Hence this remaining open case needs an argument more specific to bunkbed graphs (supposing, as we are, the conjecture holds).

In [12], the bunkbed conjecture is shown to hold true for all outerplanar graphs $G$. This result makes use of the model described below and proves a generalization of the conjecture to hypergraphs.

#### Minimal counterexamples

Let $G, x, y$ be a minimal counterexample to the bunkbed conjecture in the sense that $P_B(x_0 \leftrightarrow y_0) < P_B(x_0 \leftrightarrow y_1)$, where $B = B(G)$, but that this is not true of any minor of $G$. Recall that a minor of a graph $G$ is a graph obtained from $G$ by contracting or deleting edges. For any random graph $H$ having vertices

$x_0$, $y_0$ and $y_1$, let $\Delta = P_H(x_0 \leftrightarrow y_0) - P_H(x_0 \leftrightarrow y_1)$, and $F$ be any event. For example, $e_0\overline{e_1}$ will denote the event $\{e_0$ is open and $e_1$ is closed$\}$. Define $\Delta_H(F) := P_H(x_0 \leftrightarrow y_0|F) - P_H(x_0 \leftrightarrow y_1|F)$. Let $e$ be an edge in $G$ with both endpoints in $T$. In this case $e_0 = e_1$. Then

$$\Delta_B = p_e\Delta_{B/e} + (1-p_e)\Delta_{B-e} = p_e\Delta_{B(G/e)} + (1-p_e)\Delta_{B(G-e)},$$

which by hypothesis ($G/e$ and $G-e$ are not counterexamples) is positive. This contradicts $G$ being a counterexample. Hence there can be no edge with both endpoints in $T$.

Let $e$ be an edge in $G$ with at least one endpoint outside $T$, so that $e_0 \neq e_1$. Then

$$\Delta_B = \Delta_B(e_0e_1) + p_e(1-p_e)(\Delta_B(e_0\overline{e_1}) + \Delta_B(\overline{e_0}e_1)) + (1-p_e)^2\Delta_B(\overline{e_0e_1}).$$

By hypothesis, $\Delta_B(e_0e_1) = \Delta_{B(G/e)}$ and $\Delta_B(\overline{e_0e_1}) = \Delta_{B(G-e)}$ are both non-negative. Hence $\Delta_B(e_0\overline{e_1}) + \Delta_B(\overline{e_0}e_1)$ must be negative. Disproving this would give the desired contradiction that no minimal counterexample can exist. Hence, to show the bunkbed conjecture, it is sufficient to show that $P'(x_0 \leftrightarrow y_0) \geq P'(x_0 \leftrightarrow y_1)$, where $P'$ is the probability measure conditional on $\bigcap_{e\in E(G)}\{$ either $e_0$ is open and $e_1$ is closed or $e_0$ is closed and $e_1$ is open$\}$. Note that this model easily is thought of as a process in $G$ rather than $B(G)$.

## 4.3 Expected cluster size

To prove BBC, we would like to show that for each pair of vertices $(x, y)$, the inequality

$$P(x_0 \leftrightarrow y_0) - P(x_0 \leftrightarrow y_1) \geq 0 \tag{4.1}$$

holds.

An obvious necessary condition is

$$\sum_y (P(x_0 \leftrightarrow y_0) - P(x_0 \leftrightarrow y_1)) \geq 0, \tag{4.2}$$

for each vertex $x$.

In fact, this necessary condition is also sufficient for BBC to hold.

**Theorem 4.3.** *If* (4.2) *holds for every random graph, then* (4.1) *holds for every random graph.*

*Proof.* Let $(G, \mathbf{p})$ be a random graph. Assume for the sake of contradiction that there are two vertices $x$, $z$ in $G$ for which (4.1) fails, that is, for which $P_G(x_0 \leftrightarrow z_0) < P_G(x_0 \leftrightarrow z_1)$. For each $r \geq 0$ define a random graph $(G_r, \mathbf{p}_r)$ by adding $r$ new vertices of degree one, all connected to $z$. We leave $T$ fixed. Let the labels of the new edges be 1 and keep the labels of the original edges. A simple computation shows that

$$\sum_y (P_{G_r}(x_0 \leftrightarrow y_0) - P_{G_r}(x_0 \leftrightarrow y_1)) = \sum_y (P_G(x_0 \leftrightarrow y_0) - P_G(x_0 \leftrightarrow y_1)) + r(P_G(x_0 \leftrightarrow z_0) - P_G(x_0 \leftrightarrow z_1))$$

By hypothesis this expression tends to $-\infty$ as $r$ tends to infinity. Hence there is some $r_0$ for which this value is negative. Hence the graph $G_{r_0}$ contradicts (4.2). ☐

Actually, the left hand side of (4.2) has a more direct interpretation. To see this, observe that the random variable $|C_{x_0}|$, where $C_{x_0}$ denotes the (random) set of vertices connected to $x_0$ by open paths, can be written

$$|C_{x_0}| = \sum_{z\in B(G)} I(x_0 \leftrightarrow z),$$

where $I(x_0 \leftrightarrow z)$ is the indicator variable of the event $\{x_0 \leftrightarrow z\}$. Taking expectations, we have

$$E|C_{x_0}| = E\sum_{z\in B(G)} I(x_0 \leftrightarrow z) = \sum_{z\in B(G)} E(I(x_0 \leftrightarrow z)) = \sum_{z\in B(G)} P(x_0 \leftrightarrow z).$$

A similar argument (splitting the sum into the cases $z \in G_0$ and $z \in G_1$ (the terms corresponding to $z \in T$ are all zero)) shows that the left hand side of (4.2) may be written

$$\sum_y (P(x_0y_0) - P(x_0y_1)) = E(|C_{x_0} \cap G_0| - |C_{x_0} \cap G_1|).$$

Hence BBC is equivalent to the possibly even more intuitively clear statement that $E|C_{x_0} \cap G_0| \geq E|C_{x_0} \cap G_1|$ for each graph $G$. Observe that this is equivalent to the statement $E|C_{x_0} \cap G_0'| \geq E|C_{x_0} \cap G_1'|$.

It is now natural to sum over both $x$ and $y$ in (4.1) instead of just one of the vertices, and ask whether the result is nonnegative:

$$\sum_{x,y}(P(x_0y_0) - P(x_0y_1)) = \sum_x E(|C_{x_0} \cap G_0| - |C_{x_0} \cap G_1|) \geq 0 \qquad (4.3)$$

This is another necessary condition for (4.1), whose sufficiency for (4.1) seems harder to prove. However, it's easier to show that this statement is indeed true.

**Theorem 4.4.** *For any random graph $G$, the expression in* (4.3) *is nonnegative.*

*Proof.* Below, $Q$ denotes a partition of the vertices of $B(G)$, sums over $Q$ run over all possible such partitions, and $P(\cdot|Q)$ is the probability measure conditional on
$\mathcal{H}_Q :=$ {the equivalence classes of the relation '$x$ reaches $y$' are given by the blocks of Q}. By $P(Q)$ we mean $P(\mathcal{H}_Q)$. As usual the elements of a partition are blocks $B$, that is, pairwise disjoint subsets of vertices whose union is all of $V(B(G))$.

For a vertex $x$ in the bunkbed, let $X_x = \begin{cases} |C_x \cap G_0| - |C_x \cap G_1| & \text{if } x \in G_0' \\ 0 & \text{if } x \in T \\ |C_x \cap G_1| - |C_x \cap G_0| & \text{if } x \in G_1' \end{cases}$

Let $B$ be a block of $Q$. Since $X_x$ depend only on which vertices are reachable by open paths from which, $X_x$ is completely determined given $\mathcal{H}_Q$. To evaluate $\sum_{x\in B} E(X_x|Q)$, note that for $x \in B\cap G_0'$, we have $E(X_x|Q) = |B\cap G_0'| - |B\cap G_1'|$; for $x \in T$, $E(X_x|Q) = 0$, and for $x \in B \cap G_1'$, $E(X_x|Q) = |B \cap G_1'| - |B \cap G_0'|$. Hence

$$\sum_{x\in B} E(X_x|Q) = (|B \cap G_0'| - |B \cap G_1'|)^2.$$

Now,

$$\begin{aligned} 2\sum_{x,y\in G}(P(x_0\leftrightarrow y_0)-P(x_0\leftrightarrow y_1)) &= 2\sum_{x\in G}E(|C_{x_0}\cap G_0|-|C_{x_0}\cap G_1|) \\ &= \sum_{x\in B(G)}E(X_x) \\ &= \sum_Q\sum_{B\in Q}E(\sum_{x\in B}X_x|Q)P(Q) \\ &= \sum_Q\sum_B(|B\cap G_0'|-|B\cap G_1'|)^2P(Q)\geq 0. \end{aligned}$$

□

Note that a partial sum of the sum in (4.3) is $\sum_x(P(x_0\leftrightarrow x_0)-P(x_0\leftrightarrow x_1))=\sum_x(1-P(x_0\leftrightarrow x_1))$, which certainly is nonnegative. Do the other terms have a positive sum as well? That is, does

$$\sum_{x,y:x\neq y}P(x_0\leftrightarrow y_0)-P(x_0\leftrightarrow y_1)\geq 0 \tag{4.4}$$

hold? If so, this would show that in any bunkbed graph, there are two different vertices $x$, $y$ for which $P(x_0\leftrightarrow y_0)\geq P(x_0\leftrightarrow y_1)$. Hence the following conjecture is a weaker version of BBC.

**Conjecture 4.1.** *For any random graph $G$, (4.4) holds.*

The inequality (4.4) has an interesting reformulation:

**Claim 4.1.**

$$\sum_{x,y:x\neq y}P(x_0y_0)-P(x_0y_1)=\sum_{B\subseteq V(B(G))}\left((|B\cap G_0'|-|B\cap G_1'|)^2-(|B\cap G_0'|+|B\cap G_1'|)\right)P(B),$$

*where $P(B)$ denotes the probability of the subset $B$ of nodes forming a component in $B(G)$.*

This claim is easily proved by the method used in Theorem 4.4.

If we attempt a similar construction used in 4.3 to prove that (4.3) ⇒ (4.2), we get a non-trivial, though nonpositive, lower bound on $P(x_0\leftrightarrow y_0)-P(x_0\leftrightarrow y_1)$ in terms of $P(x_0\leftrightarrow x_1)$ and $P(y_0\leftrightarrow y_1)$.

Let $x$ and $y$ be vertices in a random graph $G$. Construct $G_r$ by joining $r-1$ new vertices each to $x$ and $y$. These new $r$-sets of vertices will be denoted by $X$ and $Y$. Probabilities taken in $G$ are denoted by $P$, those in $G_r$ by $P'$. The sum

$$\sum_{u,v\in V(G_r)}(P'(u_0\leftrightarrow v_0)-P'(u_0\leftrightarrow v_1))$$

is a quadratic polynomial in $r$ whose coefficient in front of $r^2$ necessarily is nonnegative, as we otherwise would be able to find counterexamples to (4.3). This coefficient is easily computed to be $2(P(x_0\leftrightarrow y_0)-P(x_0\leftrightarrow y_1))+(1-P(x_0\leftrightarrow x_1))+(1-P(y_0\leftrightarrow y_1))$ (these are, from the left, the contributions from the cases ($u \in X, v \in Y$ or $v \in Y, u \in X$), ($u, v \in X$), ($u, v \in Y$)). We deduce that for any graph $G$,

$$P(x_0 \leftrightarrow y_0) - P(x_0 \leftrightarrow y_1) \geq \frac{P(x_0 \leftrightarrow x_1) + P(y_0 \leftrightarrow y_1)}{2} - 1.$$

Interestingly, this inequality is an equality for graphs with $P(x_0 \leftrightarrow x_1) = P(y_0 \leftrightarrow y_1) = 1$. It seems highly probable that the left hand side is large when the right hand side is small (varying $G$).

One may prove a similar inequality concerning $E|C_{x_0} \cap G_0| - E|C_{x_0} \cap G_1|$ and $P(x_0 \leftrightarrow x_1)$ by replacing each vertex except $x$ by an $r$-clique.

## 4.4 Limiting cases

Consider a bunkbed graph with all edge parameters equal to some variable $p \in [0, 1]$.

Letting all edge parameters equal the same number is not a very strong restriction, since, by adding parallel edges and subdividing edges, any (finite) random graph can be arbitrarily well approximated (two random graphs are approximately equal if all values $P(x \leftrightarrow y)$ are approximately equal) by a random graph with all edge parameters equal to some (small) number $p$.

When $p \to 0^+$, will $P_p(x_0 \leftrightarrow y_0) - P_p(x_0 \leftrightarrow y_1)$, suitably normalised, tend to a nonnegative value?

The following asymptotics are easily verified:

$$P_p(x_0 \leftrightarrow y_0) \sim N(x_0, y_0)p^d$$

and

$$P_p(x_0 \leftrightarrow y_1) \sim N(x_0, y_1)p^{d'},$$

where $d$ is the distance between $x_0$ and $y_0$, $d'$ the distance between $x_0$ and $y_1$, and $N(s, t)$ is the number of distinct paths of minimal length between $s$ and $t$. Observe that $d \leq d'$. If $d = d'$ then $N(x_0, y_0) \geq N(x_0, y_1)$, as can be shown by mapping length $d$ paths $\alpha$ from $x_0$ to $y_1$ injectively to length $d$ paths $\beta$ from $x_0$ to $y_0$; let $\beta$ be the path following $\alpha$ to the last vertex $t$ at which $\alpha$ enters $T$. The last segment of $\beta$ will be the mirror image of the segment of $\alpha$ coming after $t$.

Graphs for which either of these inequalities are strict will hence satisfy the bunkbed conjecture in the limit $p \to 0$.

When $p \to 1$ the interesting quantities, in a sense to be specified, are $C(x_0, y_0)$ and $C(x_0, y_1)$, where $C(s, t)$ is the size of a minimum cut separating $s$ and $t$. Suppose the number of edges in the bunkbed is $m$, and define

$$A_k = \#\{\text{subgraphs with } x_0 \leftrightarrow y_0 \text{ and exactly } m - k \text{ edges}\},$$

and

$$B_k = \#\{\text{subgraphs with } x_0 \leftrightarrow y_1 \text{ and exactly } m - k \text{ edges}\}.$$

Then $P(x_0 \leftrightarrow y_0) = \sum_k A_k p^{m-k}(1-p)^k$ and $P(x_0 \leftrightarrow y_1) = \sum_k B_k p^{m-k}(1-p)^k$.

Note that if $k < \min C(x_0, y_0), C(x_0, y_1)) =: D$ then $A_k = \binom{m}{k} = B_k$. Hence $f(p) := P_p(x_0 \leftrightarrow y_0) - P_p(x_0 \leftrightarrow y_1)$ satisfies $f(p) = \alpha(1-p)^D + o((1-p)^D)$ for $p$ close to 1 and some constant $\alpha$. For the bunkbed conjecture to hold for $p$ arbitrarily close to 1, we must have $\alpha \geq 0$. But $\alpha = C(x_0, y_0) - C(x_0, y_1)$ whence $C(x_0, y_0) \geq C(x_0, y_1)$. By the min cut-max flow theorem this is the same as $F(x_0, y_0) \geq F(x_0, y_1)$, where $F(s,t)$ is the maximum flow between $s$ and $t$ in the flow network obtained by turning each edge of the bunkbed into a two-way unit capacity edge. Thus, in the limit $p \to 1^-$, the bunkbed conjecture predicts

$$F(x_0, y_0) \geq F(x_0, y_1).$$

Indeed, with some effort, we can show this to hold true in any bunkbed flow network.

**Theorem 4.5.** *In the notation above,*

$$F(x_0, y_0) \geq F(x_0, y_1).$$

*Proof.* Let $F$ be a maximal flow from $x_0$ to $y_1$; that is, $x_0$ is the unique source of $F$, $y_1$ the unique sink. A flow is said to be feasible if it satisfies the capacity constraints at all edges, and a maximal flow is understood to be feasible by definitio by definition.. Since our capacity constraints are all integers (indeed, all 1), by the well-known integrity theorem from flow theory (see, for example, theorem 8.1 [7]), we can assume $F$ to be integer valued. Since the capacity constraint on each edge is 1, this means that $F$ takes either the value 0 or 1 on each edge.

We prove the theorem by finding a feasible flow $f'$ from $x_0$ to $y_0$ of same strength as $F$.

Observe that $F$ can be decomposed as $F = F_0 + F_T + F_1$, where $F_0$ is zero on all edges having no endpoint in $G'_0$, $F_T$ is zero on edges not having both endpoints in $T$, and $F_1$ is zero on all edges having no endpoint in $G'_1$.

We can think of $F_0$ as a flow from $x_0$ to $T$, $F_T$ a flow inside $T$, and $F_1$ a flow from $T$ to $y_1$.

Let $f = F_0 + F_T + \varphi F_1$, where $\varphi F_1$ is the flow satisfying $\varphi F_1(u, v) = F_1(\varphi u, \varphi v)$ for all $u, v$. Here $\varphi$ is the involution turning the bunkbed 'up-side down', that is, $\varphi(z_0) = z_1$, $\varphi(z_1) = z_0$, and $\varphi(t) = t$ for all $z \in V(G)$, $t \in T$.

Observe that $f$ is zero on all edges having at least one endpoint in $G'_1$.

The flow $f$ is clearly a flow from $x_0$ to $y_0$ of the same strength as $F$.

However $f$ need not be a feasible flow, since the flow of $f$ over some edges might equal 2; we call such edges *$f$-bad*. It is easy to see that if the flow of $f$ is not 2 over any edge, then $f$ is feasible.

Given a directed cycle or directed path $X$, denote by $u(X)$ the unit flow associated to $X$ given by sending one unit of flow along each edge of $X$.

Since the only source of $F_0$ is $x_0$ and all sinks lie in $T$, $F_0$ can be written as the sum of $u(P)$, where the $P$ are directed paths going from $x_0$ to vertices in $T$. These paths will be called *U paths* ('U' as in going Up from $x_0$ to $T$). Similarly, the paths going from $T$ to $y_0$ corresponding to a unit flow decomposition of $\varphi F_1$ will be called *D paths*.

Clearly, any two U paths and any two D paths are edge disjoint.

We would like to turn $f$ into a feasible flow $f'$ of the same strength.

To achieve this, we will first find a collection $\mathcal{C}$ of directed paths and directed cycles satisfying certain properties.

Initially, our collection will consist of directed paths only. These paths will be initial segments of D paths. Specifically, for each D path $\alpha$ containing at least one $f$-bad edge, include in $\mathcal{C}$ the initial segment of $\alpha$ ending with the last $f$-bad edge occurring along $\alpha$.

Observe that $\mathcal{C}$ satisfies the following properties

- Each path $P$ in $\mathcal{C}$ begins with an edge incident to some vertex in $T$ and ends with an $f$-bad edge. We denote this $f$-bad edge by $e(P)$.
- No two elements in $\mathcal{C}$ share an edge in the same direction.
- The set of constraints violated by $f - u(X)$ is a subset of those violated by $X$, for each element (path or cycle) $X$ in $\mathcal{C}$. (Phrased differently, the flow of $f$ is at least 0 at any edge of $X$, taking the positive direction to be the direction of $X$.)
- Each $f$-bad edge is included in exactly one element of $\mathcal{C}$ (in the direction of flow of $e$).

Below we will define an operation on $\mathcal{C}$ which preserves these properties.

Note that each $f$-bad edge is included in exactly one D path and exactly one U path. Conversely, any edge which is shared in the same direction by one U path and one D path is $f$-bad.

Let $P$ be a path in $\mathcal{C}$, $e = e(P)$ its final (necessarily $f$-bad) edge. Denote by $S(P)$ the segment coming after $e$ of the unique U path going through $e$. Hence $S(P)$ is a path starting with an edge incident to $e(P)$ and ending with an edge incident to some vertex of $T$ (the sink of the corresponding U path; note that the direction of flow of the U path must coincide with the direction of flow of $e$).

When the operation cannot be applied any further, the resulting collection will have the following additional property

- For each path $P$, there are no $f$-bad edges along $S(P)$.

Later, in order to show that the algorithm below terminates after a finite number of operations, we will show that the (initially finite!) integer describing the number of *jumps*, to be defined, decreases strictly after each operation. Each element of $\mathcal{C}$ (path or cycle) consists of $f$-bad edges and subpaths (which include no $f$-bad edges themselves) joining these. If the element is a path, in addition we have an initial segment starting at a vertex of $T$ and ending at the first $f$-bad edge occurring along this path.

We call such subpaths *steps* (hence elements of $\mathcal{C}$ are alternating concatenations of $f$-bad edges and steps). Of course, a step might have zero length, corresponding to two consecutive $f$-bad edges. At each point we will consider each step of each element in $C$ to be a *jump* or not. In our initial collection we consider all steps to be jumps.

Below, the interesting quantity will be the total number of jumps among all elements of $\mathcal{C}$. This integer will be called the *jump number* (of $C$).

We will use the fact (which is a consequence of how our initial collection and the operation below are defined) that each step of any element in $\mathcal{C}$ that is not

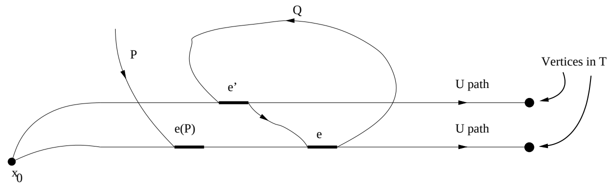


Figure 4.5: Case (i): before

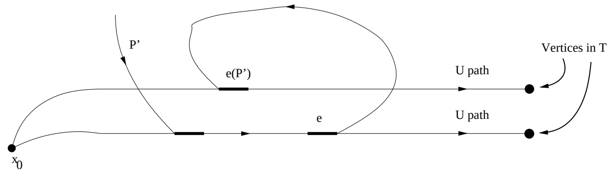


Figure 4.6: Case (i): after

a jump, is a subpath of a U path, and that the two $f$-bad edges this step links lie on this same U path (and the flow of these edges are in the direction of the U path).

## The operation

Suppose $P$ is a path in $C$ for which $S(P)$ includes at least one $f$-bad edge. Let $e$ be the first $f$-bad edge occurring along $S(P)$. Let $Q$ be the (unique) element of $\mathcal{C}$ traversing $e$. A few cases arise. Cf. figures 4.5 - 4.10. These figures convey the general picture rather well, bearing in mind the following remarks: two $f$-bad edges might be consecutive, $e'$ (defined separately in each case) might lie on the same U path as $e$ (but necessarily before $e(P)$ on this U path, by choice of $e$), and in case (ii), $Q$ might end at $e$.

### Case (i). $Q$ is a cycle

As will be seen in case (iii) (the only point where cycles are added to $\mathcal{C}$), any cycle in $\mathcal{C}$ will have at least two $f$-bad edges. Let $e'$ be the $f$-bad edge occurring before $e$ on $Q$.

Let $P'$ be the path following $P$ to $e(P)$, taking $S(P)$ to $e$, and ending with the part of $Q$ between $e$ and $e'$ (in the direction of $Q$). This will include all $f$-bad edges of $Q$ in $P'$. The jumps of $P'$ are defined to be the corresponding jumps of $P$ and $Q$. In particular, the step of $P'$ from $e(P)$ to $e$ is not considered a jump of $P'$.

Remove $P$, $Q$ and add $P'$ to $\mathcal{C}$.

Since the step of $Q$ between $e'$ and $e$ is a jump (this follows from our fact concerning jumps stated above; if this step was not a jump, $e'$ would be an $f$-bad edge strictly between $e(P)$ and $e$ on $S(P)$, contradicting the choice of $e$), the jump number decreases by 1 in this case.

### Case (ii). $Q$ is a path different from $P$

Suppose there is no $f$-bad edge occurring along $Q$ before $e$. In this case simply remove $Q$ from $\mathcal{C}$.

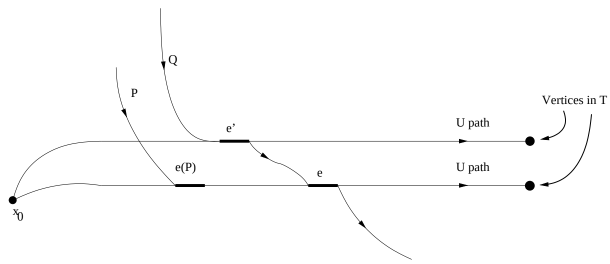


Figure 4.7: Case (ii): before

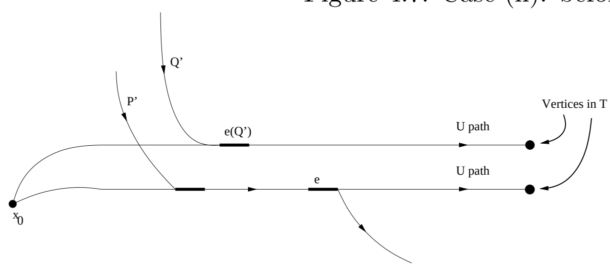


Figure 4.8: Case (ii): after

Otherwise, let $e'$ be the $f$-bad edge along $Q$ before $e$. Let $Q'$ be $Q$ cut off after $e'$, where a step of $Q'$ is considered a jump if and only if the corresponding step in $Q$ is a jump. Remove $Q$ and add $Q'$ to $\mathcal{C}$ in this case.

Let $P'$ be the path starting with $P$, following $S(P)$ to $e$, ending with part of $Q$ coming after $e$ (of course, this part might be empty). The jumps of $P'$ are defined to be the corresponding jumps of $P$ and $Q$. In particular the step of $P'$ between $e(P)$ and $e$ is not considered a jump of $P'$. Add $P'$ to $\mathcal{C}$.

In this case the jump number decreases by one (we 'lose' the step, which is a jump, of $Q$ from $e'$ to $e$. This step being a jump is motivated in the same way as in (i)).

**Case (iii).** $Q = P$

If there is some $f$-bad edge occurring along $P$ before $e$, then let $R$ be $P$ cut off after the latest such edge (we define the jumps of $R$ to be precisely those steps of $R$ that correspond to jumps of $P$). Add $R$ to $\mathcal{C}$ in this case.

Let $O$ be the cycle 'starting' at $e$, following $P$ to $e(P)$, then following $S(P)$ back to $e$. The jumps of $O$ are the same as those of $P$ from $e$ to $e(P)$. In particular the step of $O$ from $e(P)$ to $e$ is not considered a jump. Add $O$ and remove $P$ from $\mathcal{C}$.

The jump number decreases by 1 in this case as well, since the step of $P$ between $e'$ and $e$ is a jump.

Note that in each case of the operation, the jump number decreases by 1. It follows that the operation can be applied only a finite number of times.

Now let $\mathcal{C}$ be the collection obtained by repeatedly applying the operation to the initial collection until this can be done no more.

Given a path $P$ in $\mathcal{C}$, let $E(P)$ be the path $P$ followed by $S(P)$. As a

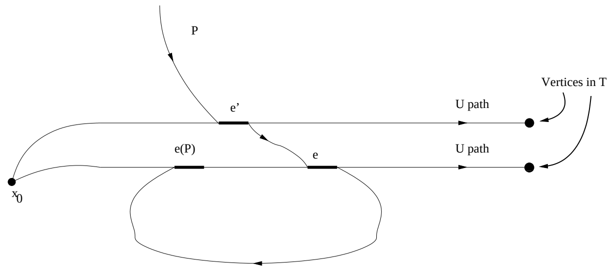


Figure 4.9: Case (iii): before

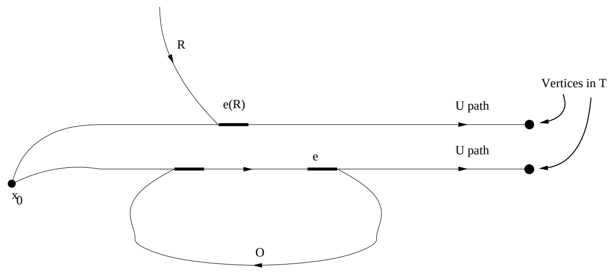


Figure 4.10: Case (iii): after

consequence of the operation not being applicable to $\mathcal{C}$, no two $E(P)$ share an edge in the same direction (and, being subpaths of U paths, the $S(P)$ are all edge disjoint). In addition, no $E(P)$ shares an edge in the same direction with any cycle in $\mathcal{C}$.

We claim that

$$f' = f + \sum_{P \in \mathcal{C} \text{ is a path}} (\varphi u(E(P)) - u(E(P))) - \sum_{O \in \mathcal{C} \text{ is a cycle}} u(O)$$

is a feasible flow from $x_0$ to $y_0$ of the same strength as $F$.

Clearly, the strengths of the source at $x_0$ and the sink at $y_0$ are of the same strength as $F$.

There are no other sources or sinks in $f'$, as subtracting circulation flows or adding flows of the form $u(P) - \varphi u(P)$, where $\varphi$ is any function of flows fixing the endpoints of the directed path $P$ (the endpoints of our $P$:s are in $T$) does not introduce sources or sinks.

Finally we need to show that $f'$ violates no constraints.

For edges with at least one endpoint in $G_1'$, this is a consequence of the elements of $\mathcal{C}$ sharing no edge in the same direction, and it is immediate for edges with both endpoints in $T$, as $f'$ conincides with $f$ on such edges.

By the properties of $\mathcal{C}$, decreasing the flow along any element of $C$ does not violate constraints already violated. Moreover, the flow over any $f$-bad edge along an element of $\mathcal{C}$ is decreased by 1. Since each $f$-bad edge is included in exactly one $\mathcal{C}$-element, $f'$ has no '$f'$-bad' edges.

This proves our claim, proving the theorem. □

We now consider an example, illustrating the idea of the proof. Consider figure 4.11. The horizontal lines are the U paths containing at least one $f$-bad

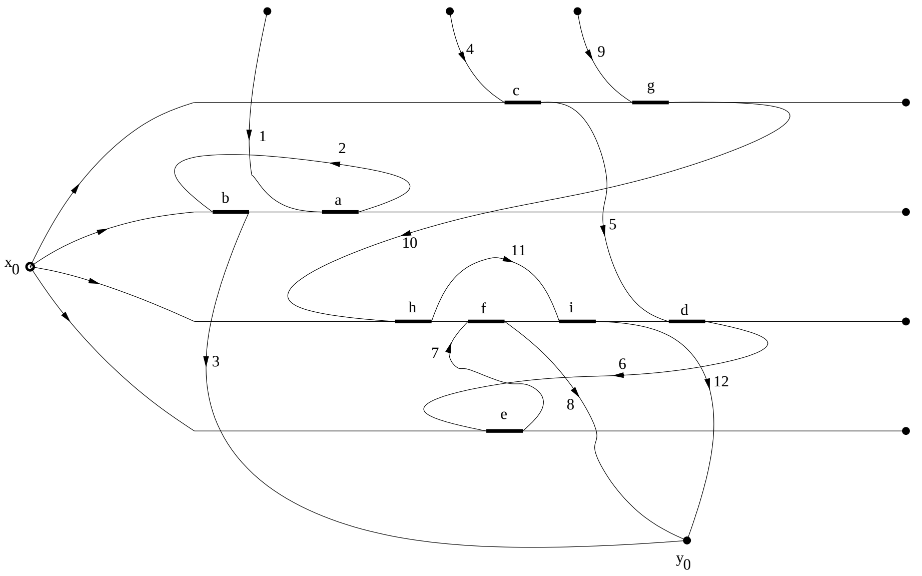


Figure 4.11: An example

edge. The D paths containing at least one $f$-bad edge are $1a2b3$, $4c5d6e7f8$, and $9g10h11i12$. Here the numbers refer to maximal subpaths of D paths containing no $f$-bad edges (that is, steps, if we were talking about paths from $\mathcal{C}$; we are not yet). The $f$-bad edges are named $a, \ldots, i$ and are drawn bold. There will be no ambiguity between the flow $f$ and the $f$-bad edge $f$. The vertices indicated with black dots, apart from $x_0$ and $y_0$, are vertices in $T$. Initially we define $\mathcal{C}$ to be $\{1a2b, 4c5d6e7f, 9g10h11i\}$.

Applying the operation (case (ii)) to the $f$-bad edge i, we get

$$\mathcal{C} = \{1a2b, 4c5d6e7f \to i, 9g10h\},$$

where we use the notation $f \to i$ to indicate following the common U path of $f$ and $i$ from $f$ to $i$.

Next we apply the operation (case (iii)) to the $f$-bad edge d, and get

$$\mathcal{C} = \{1a2b, 4c, 6e7f \to i \to d, 9g10h, 4c\},$$

where $6e7f \to i \to d$ is a cycle (which we could just as well have written $e7f \to i \to d6$ et.c.).

Apply the operation (case (ii)) to $g$, yielding

$$\mathcal{C} = \{1a2b, 4c \to g10h, 6e7f \to i \to d\}.$$

Operate (case (i)) on $f$ to get

$$\mathcal{C} = \{1a2b, 4c \to g10h \to f \to i \to d6e\}$$

Finally, operate (case (ii)) on $r$, to give the final collection

$$\mathcal{C} = \{2b \to a, 4c \to g10h \to f \to i \to d6e\}.$$

Above steps correspond to numbered subpaths and arrows ('→'). The steps considered to be jumps are precisely the numbered ones (omitting, of course, $3, 8$, and 12). Hence, the jump number of $\mathcal{C}$ given a description of $\mathcal{C}$ as above is simply number of numbers occurring in this description.

Observe that $C$ contains one path, $P$, and one cycle, $O$. We construct $f'$ from $f$ by decreasing the flow of $f$ along $O$ and 'flipping' $E(P)$ from $G_0$ to $G_1$. Here $E(P)$ is the path $P$ followed by $S(P)$, as before.

In the proof above, we could actually start with any feasible integer valued (not necessarily maximal) flow from $x_0$ to $y_1$. Hence the proof actually describes a way to map families of disjoint $x_0y_1$ paths into families of disjoint $x_0y_0$ paths (of the same number), which might be a useful tool in proving the full bunkbed conjecture.

## 4.5 When the transversal set is a cut set

As observed in [12], when $x$ and $y$ belong to different components of $G - T$, $P(x_0 \leftrightarrow y_0) = P(x_0 \leftrightarrow y_1)$, as the following mirror argument shows (taken from [12]).

**Lemma 4.1.** *If $x$ and $y$ belong to different components of $G-T$, then $P(x_0 \leftrightarrow y_0) = P(x_0 \leftrightarrow y_1)$.*

*Proof.* We give an involution $\varphi$ on the probability space $B(G)$ taking $\{x_0 \leftrightarrow y_0\}$ to $\{x_0 \leftrightarrow y_1\}$. This shows that the probability of the these two events coincide. Edges in $B(G)$ are called *upper*, *lower*, or *central* if they have at least one endpoint in $G_0$, at least one in $G_1$ or both in $T$ respectively. The event $\varphi(\omega)$ is obtained from $\omega$ by interchanging the state of upper and lower edges having at least one endpoint whose image in $G$ is in the component of $y$ in $G - T$. The states of all other edges (in particular, of all central edges) are kept constant. It is readily verfied that $\varphi$ satisfies the properties sought for. □

The converse of the above lemma does not hold, as may be seen by choosing any graph not satisfying the hypothesis but satisfying $P(x_0 \leftrightarrow x_1) = 1$ or $P(y_0 \leftrightarrow y_1) = 1$. However, the converse seems reasonable to hold if we in addition assume all entries of $\mathbf{p}$ to be strictly less than 1. Note that this does not rule out important cases, since, by contracting edges, a bunkbed graph with some entry $p_e = 1$ is equivalent to a smaller bunkbed graph. A similar remark applies to the case $p_e = 0$, although this is not needed here.

We formulate the converse as a conjecture.

**Conjecture 4.2.** *Suppose $G$ is a random graph having no edge label equal to 1. Let $x$ and $y$ be two vertices of $G$ with $P_{B(G)}(x_0 \leftrightarrow y_0) = P_{B(G)}(x_0 \leftrightarrow y_1)$. Then any path between $x$ and $y$ passes through $T$.*

Proving this conjecture is not as peripheral as might seem at first sight.

**Theorem 4.6.** *Conjecture 4.2 implies the bunkbed conjecture*

*Proof.* Suppose Conjecture 4.2 is true. Let $G$ be a counterexample to the bunkbed conjecture with $P_{\mathbf{p_0}}(x_0y_0) < P_{\mathbf{p_0}}(x_0y_1)$ for some vertices $x$, $y$ in $G$ and some probability vector $\mathbf{p_0}$.

Define $f : [0,1]^{E(G)} \to \mathbb{R}$ by $f(\mathbf{p}) = P_{\mathbf{p}}(x_0y_0) - P_{\mathbf{p}}(x_0y_1)$ Since $P(x_0y_0) \neq P(x_0y_1)$, by Lemma 4.1, there is a path from $x$ to $y$ contained in $G-T$. Choosing a probability vector $\mathbf{p}'$ which is equal to 1 on the edges of this path and 0 elsewhere, there is some point in $[0,1]^{E(G)}$ for which $f$ is positive. Since $f(\mathbf{p_0}) < 0$ and $f$ is continuous (indeed, a polynomial in $\mathbf{p}$), the mean value theorem (observing that $[0,1]^{E(G)}$ is connected!) gives that there is some point $\xi \in (0,1)^{E(G)}$ for which $f(\xi) = 0$. By the converse of Claim 1, this means that $x$ is separated from $y$ in $G - T$, which is a contradiction. Hence there can be no such counterexample $G$. □

Let $A_k = \#\{\text{configurations in which } x_0 \leftrightarrow y_0 \text{ and the number of open edges is } k\}$, $B_k = \#\{\text{configurations in which } x_0 \leftrightarrow y_1 \text{ and the number of open edges is } k\}$.

Denote the total number of edges in the graph by $m$.

We have

$$P(x_0 \leftrightarrow y_0) = \sum_{k=1}^{m} A_k p^k (1-p)^{m-k}$$

and

$$P(x_0 \leftrightarrow y_1) = \sum_{k=1}^{m} B_k p^k (1-p)^{m-k}.$$

If $p$ is transcendent over the reals, we must have $A_k = B_k$ for each $k$ (observing that $\left(p \mapsto p^k(1-p)^{m-k}\right)_0^m$ constitute a basis of the real vector space of all polynomials of degree $\leq m$), meaning that there is an involution $\varphi$ satisfying the requirements in the proof of Lemma 4.1. This is a first step in boldly reversing the proof of lemma 4.1. It does however seem hard to extend this idea to algebraic $p$, let alone to deduce that $T$ separates $x$ and $y$ from the fact that there is such an involution $\varphi$.

## 4.6 Differentiation

Viewing $f_B(\mathbf{p}) := P_{\mathbf{p}}(x_0 \leftrightarrow y_0) - P_{\mathbf{p}}(x_0 \leftrightarrow y_1)$ as a function of a variable $\mathbf{p} \in ]0,1[^{E(G)}$, we can derive some conditions that a minimal counterexample to the bunkbed conjecture must satisfy.

Let $(G, \mathbf{p})$ be a counterexample to the bunkbed conjecture such that for any edge $e$ in $G$, neither $G/e$ nor $G - e$ is a counterexample. Let $B = B(G)$.

If $H$ is a sub-random graph of $B$ containing $x_0, y_0, x_1, y_1$, then $f_H$ will denote $P_H(x_0 \leftrightarrow y_0) - P_H(x_0 \leftrightarrow y_1)$.

Note that $f_B = p_e^2 f_{B/e_0,e_1} + p_e(1-p_e) f_{B/e_0-e_1} + p_e(1-p_e) f_{B/e_1-e_0} + (1-p_e)^2 f_{B-e_0,e_1}$. Since $G$ is minimal, $f$ must be positive for $p_e = 0$ and $p_e = 1$, as these cases may be modeled by smaller bunkbed graphs. Moreover, for $f$ to be negative for some value of $p_e$ in $]0,1[$, the coefficient in front of $p_e^2$ in $f$ must be positive. Note that $B/e_0, e_1 = B(G/e)$. If $p_m$ is the value of $p$ mimimizing $f(p)$ we must have $0 < p_m < 1$.

The conditions $\frac{d^2 f_B}{dp_e^2} > 0$ and $p_m < 1$ reduce to

$$f_{B(G/e)} + f_{B(G-e)} > f_{B/e_0-e_1} + f_{B-e_0/e_1}$$

and

$$f_{B(G/e)} > \frac{f_{B/e_0-e_1} + f_{B-e_0/e_1}}{2}.$$

## 4.7 Electrical networks

The bunkbed conjecture has an interesting, and much easier-to-prove, analogue for electrical networks, as follows.

**Theorem 4.7.** *Let $G$ be a graph, where each edge is assigned a positive real number, called the resistance of that edge. Construct a bunkbed network in the obvious way, having chosen some transversal set $T$ of nodes from $G$. Suppose the electric potential $V$ satisfies the boundary values $V(x_0) = 0$, $V(y_0) = V(y_1) = 1$. Then the electric current entering $y_0$ is at least as large as the electric current entering $y_1$.*

*Proof.* For any vertex $x$ in the (bunkbed) network, let $I(x)$ denote the current leaving at $x$, and $I(x, y) = I(y) - I(x)$.

It is sufficient to show that for any vertex $u \in G$, $V(u_0) \leq V(u_1)$. From this it will follow that

$$I(x_0, y_0)-I(x_0, y_1) = \sum_{u\in N(y)} \frac{V(y_0) - V(u_0) - (V(y_1) - V(u_1))}{R_{uv}} = \sum_{u\in N(y)} \frac{V(u_1) - V(u_0)}{R_{uv}} \geq 0.$$

Let $\gamma(u) = V(u_1) - V(u_0)$ for $u \in V(G)$. This gives an electrical potential $\gamma$ defined on $G$ (rather than $B(G)$), and our task is to show that $\gamma$ is a nonnegative function.

By Kirchoff's law, for any vertex $u \notin \{x_0, y_0, y_1\}$, we have $I(u) = 0$, and hence $V(u) = \left(\sum_{v\in N(u)} \frac{1}{R_{uv}}\right)^{-1} \sum_{v\in N(u)} \frac{V(v)}{R_{uv}}$. Hence the electric potential of $u$ is a mean value of the electric potentials at the neighbouring vertices of $u$ (in the sense that $\min_{v\in N(u)} V(v) \leq V(u) \leq \max_{v\in N(u)} V(v)$ holds). By a standard argument, we deduce the maximum principle; min{ boundary values} $\leq V(u) \leq$ max{ boundary values} for each vertex $u$. By applying the maximum principle to $V$ in the bunkbed graph we obtain $V(x_1) \geq 0$, since all boundary values are nonnegative (indeed, they are 0 and 1).

Now $\gamma$ has the boundary value $\gamma(y) = 0$. Note that we may also assume $\gamma(y) = V(x_1)$ to be fixed (supplying part of the solution $V$ with the original boundary values as a new boundary value can not change the new solution $V$).

By applying the maximum principle to $\gamma$ in $G$, we obtain $\gamma(u) \geq 0$ for every vertex $u$. □

Note that $T$ separating $x$ and $y$ in $G$ is the same thing as saying that the component of $y$ in $G - T$ is joined to a grounded boundary. Hence $\gamma = 0$ in this component, meaning that $I(y_0) = I(y_1)$ in this case. This statement should be compared with the case of equality in the (original) bunkbed conjecture.

## Electrical networks and probability

In [2] it is proved that $R(x_0, y_0) \leq R(x_0, y_1)$ for any bunkbed graph, where $R(s, t)$ is the *effective resistance between* $s$ and $t$ in the network. The effective resistance could be defined by solving $1 = R(s, t) \cdot I$, where $I$ is the strength of the electrical current flowing from $s$ to $t$ when fixing the potential of $s$ and $t$ to 1 and 0 respectively. This result can be interpreted in terms of random walks on the bunkbed; a random walk starting at $x_0$ in each step moving to a neighbour chosen uniformly at random is more likely to reach $y_0$ before it reaches $y_1$ than to reach $y_1$ before $y_0$.

Two resistors of strengths $R_1$, $R_2$ connected in series have a combined resistance $R_{ser} = R_1 + R_2$. When connecting two resistors in parallel, the combined resistance is $R_{par} = \frac{R_1+R_2}{R_1R_2} = \frac{R_{ser}}{R_1R_2}$. The corresponding values for random graphs are $p_{ser} = p_1p_2$ and $p_{par} = p_1 + p_2 - p_1p_2 = p_1 + p_2 - p_{ser}$.

Now we have an interesting symmetry:

$$\begin{array}{rr} \ln p_{ser} = \ln p_1 + \ln p_2 & R_{ser} = R_1 + R_2 \\ p_{par} = p_1 + p_2 - p_{ser} & \ln R_{par} = \ln R_{ser} - \ln R_1 - \ln R_2 \end{array}$$

# Chapter 5

# Related topics

## 5.1 Potentials and random graphs

In the previous section, we saw that the introduction of a potential simplified the calculation of the analogue of reachability probabilities. Hence one would like to have a similar notion in a random graph. This can be achieved as follows. Let $x_0$ denote some distinguished vertex in $G$. Associate each edge $e$ with a random variable $U_e$ taking values in $[0, 1]$. Let $p_e = \int_0^1 P(U_e < q)dq$. Assume $(U_e)_{e \in E}$ to be independent, though not necessarily identically distributed.

In the other direction, given $p_e$ for each edge $e$ of $B(G)$, there are many choices of $(U_e)$. Actually, allowing $U_{e_0}$ and $U_{e_1}$ to be dependent gives another interesting model, since the BBC is trivial in the case where $U_{e_0}$ are $U_{e_1}$ equal.

The sought-for potential is $\nu(z) := E(\inf\{q :$ the edges with $U_e < q$ connect $x_0$ and $z\})$.

We recover the original percolation model by choosing a $q \in [0, 1]$ randomly and uniformly, declaring an edge $e$ to be open if (and only if) $U_e < q$, and the bunkbed conjecture amounts to the statement $\nu(y_0) \leq \nu(y_1)$.

The potential $\nu$ does not, however, satisfy any relations analogous to the electrical potential $V$ in the previous section.

## 5.2 The random cluster model

**Definition 5.1.** *The random cluster measure $P$, with parameter $q$, on a graph $G$ where each edge $e$ is labeled by some real number $p_e \in [0, 1]$ is defined by*

$$P(\omega) \propto q^{c(\omega)} \prod_{e:\omega_e=1} p_e \prod_{e:\omega_e=0} (1 - p_e),$$

*for subgraphs $\omega$ of $G$. Here $c(\omega)$ is the number of connected components in $\omega$, and, as before, $\omega_e$ is the indicator variable of the edge $e \in E(G)$ being included in $E(\omega)$.*

For $q = 1$, the random cluster measure is just the random graph measure defined earlier. When $q \neq 1$, edges will not be open independently of each other.

For $q = 2$, the corresponding bunkbed conjecture is true, shown in [8]. This result uses to the strong connection between the well studied Ising model from statistical physics and the random cluster model with $q = 2$.

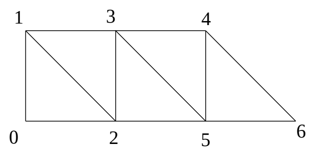


Figure 5.1: A random graph


Figure 5.2: Replacing each edge with two identical edges in parallel in this graph changes the $x$-order.

It is interesting to note that when all edge labels $p_e$ equal the same number $p$, the limit as $q \to 0$ and $q/p \to 0$ of the random cluster measure is the uniform measure on all spanning trees of $G$, assuming $G$ is connected (see [5]).

## 5.3 Ordering $V(G)$

To prove BBC, one could try to reduce counterexamples to more 'extreme' counterexamples. For example, reducing to counterexamples where all edge labels are small would at the least give a proof of BBC for graphs in which there is a $x_0y_0$-path whose probability of being open is strictly greater than that of any $x_0y_1$-path.

Let $G$ be a percolation graph with some distinguished vertex $x$, called the 'root'. Generically (varying $\mathbf{p} \in [0,1]^{|E|}$), the vertices $y$ are totally ordered by the rank function $P(x \leftrightarrow y)$ (for definiteness, let $y \leq_x z \Leftrightarrow P(x \leftrightarrow y) \geq P(x \leftrightarrow z)$). A total order on the vertices of $G$ may be identified with a permutation $\sigma \in S_{|V(G)|}$, which we refer to as the *percolation order from $x$ in $G$*, or just *$x$-order*.

An interesting question is how the $x$-order of a graph changes when changing $\mathbf{p}$. Note that the relation '$\mathbf{p}$ and $\mathbf{p'}$ give the same $x$-order' partitions almost all of $[0,1]^{|E|}$ into components $C_\sigma$.

Some natural questions are

- If $C_\sigma \neq \emptyset$, does the closure $\bar{C_\sigma}$ include 0?
- Is $C_\sigma$ connected as a subset of $\mathbb{R}^{|E(G)|}$?

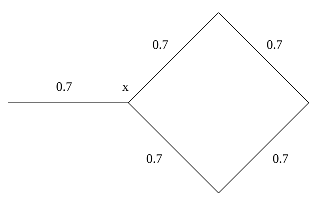


Figure 5.3: Replacing each edge with two identical edges in series in this graph changes the $x$-order.

- What properties does the set $R_G := \{\sigma : C_\sigma \neq \emptyset\}$ have?
- Substitute each edge $e$ of $G$ with two parallel (independent) copies of $e$. Does the percolation order (from any vertex) change?
- Substitute each edge $e$ of $G$ with two (independent) copies of $e$ in series. Does the percolation order change?

A positive answer to the first two questions would make a corresponding statement for bunkbed graphs (that is, varying the probabilties in $G$ but ordering the vertices in all of $B(G)$) reasonable to hold as well. This would enable us to transfer any counterexample to the bunkbed conjecture to another counterexample with all edge labels being less than some arbitrarily small chosen constant.

The answers to the last two questions are both 'in general, yes'; counterexamples are given in Figures 5.3 and 5.3.

### An example

If we vary the labels $p_1$ of all edges different from $\{2,5\}$ and the label $p_2$ of the edge $\{2,5\}$ in the graph in Figure 5.3, only two $x$-orders arise for $x = 0$: (0215346) and (0213546).

# Bibliography

[1] Béla Bóllobás, Random graphs, Cambridge University Press (2001).

[2] Béla Bollobás and Graham Brightwell, Random walks and electrical resistances in products of graphs, *Disc. Applied Math.* **73** (1997), 69-79.

[3] Colin Mc Diarmid, General Percolation and random processes, Adv. Appl. Prob. 13, 40-60 (1981).

[4] Colin McDiarmid, General percolation and oriented matroids, Annals Discrete Math., **33**, 187-197 (1987).

[5] G. R. Grimmett and S. N. Winkler, Negative association in uniform forests and connected graphs, Random Structures and Algorithms **24**, 444-460 (2004).

[6] Jacob van den Berg and Jeff Kahn, A correlation inequality for connection events in percolation, Annals of Probability **29** No. 1 (2001), 123-126.

[7] L.R. Ford, Jr. and D.R. Fulkerson, Flows in Networks, Princeton University Press (1962).

[8] Olle Häggström, Probability on Bunkbed Graphs, *Proceedings of FPSAC'03, Formal Power Series and Algebraic Combinatorics* Linköping, Sweden 2003, available at www.fpsac.org/FPSAC03/ARTICLES/42.pdf.

[9] Sven Erick Alm and Svante Linusson, A counter-intuitive correlation in a random tournament (2009). To appear in Comb. Prob. and Computing.

[10] Sven Erick Alm and Svante Linusson, Correlations for paths in random orientations of $G(n, p)$ (2009), arXiv:0906.0720.

[11] Svante Linusson, A note on correlations in randomly oriented graphs, *Preprint 2009.* arXiv:0905.2881

[12] Svante Linusson, On percolation and the bunkbed conjecture (2009). To appear in Comb. Prob. and Computing.

[13] T. E. Harris, A lower bound for the critical probability in a certain percolation process, *Proc. Cambridge Philos. Soc.* **56** 13-20 (1960).

LINKÖPING UNIVERSITY
ELECTRONIC PRESS

LINKÖPINGS UNIVERSITET